Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 3090043, 13 pages
https://doi.org/10.1155/2022/3090043

Research Article

Investigation on the Identity Construction of Young Foreign Language Teachers in Colleges and Universities Based on Feature Selection Algorithm

Lingzhi Yao

Guangzhou Huashang College, Guangzhou, Guangdong 511300, China

Correspondence should be addressed to Lingzhi Yao; yaolingzhi@gdhsc.edu.cn

Received 13 May 2022; Accepted 23 June 2022; Published 30 July 2022

Academic Editor: Xiantao Jiang

With the gradual implementation of the strategic goal of strengthening the country in education, people's attention to students' learning is also increasing day by day. At the same time, the teaching effect of individual teachers depends to a large extent on the identity of teachers themselves, so it is necessary to improve the construction of teachers' identity. Most of the research on teachers' identity construction starts from the theory of identity, which does not play a significant role in the critical application in practical teaching activities. In response to these problems, this paper will use the ReliefF algorithm and the Spark algorithm in the feature selection algorithm to scientifically process the identity construction, and implement the application of the improved ReliefF feature selection algorithm and the feature selection application steps based on the Spark algorithm respectively. The experimental results show that the improved ReliefF algorithm has better feature selection accuracy when the feature range is 30% to 40%. It shows that the construction of teacher identity based on feature selection algorithm can provide an objective basis for the realization of teacher identity.

1. Introduction

With the deepening of reform and opening up and the increasing emphasis on education and teaching, more and more people are paying attention to education and teaching. The reform of traditional education is also gradually being carried out in this environment. The purpose is to improve the learning efficiency of students in the classroom and the teaching efficiency of teachers. In this whole, in addition to the change of students' learning methods, the discussion of the integrity of teachers' own identity has also become a major breakthrough in the upgrading of classroom education. Teachers' construction of their own identity can deepen their understanding of individual values, which are a sense of identity for life. At the same time, the perfection of the teacher's identity construction, for the individual of this profession not only provides the internal development motivation but also requires the teacher himself to have the perception of the identity of this profession. It is necessary to achieve the goal of recognition of the identity of the teacher by others. In the discussion, there is no obvious correlation between the identity construction of teachers and the specific methods used by teachers themselves, but more of a sound psychological state developed by teachers themselves. The importance of teachers' identity construction is thus highlighted. In order to conduct scientific research on this phenomenon, this paper will discuss the feature selection algorithm.

2. Related Work

Regarding the identity construction of teachers, due to the great attention paid to education and teaching in recent years and the increasing demands of the public for classroom teaching, different scholars have made various corresponding studies. The research done by scholars C. Zhang and Y. Zhang was to avoid the need for native language teachers and to better understand the identity of foreign teachers [1]. Zacharias conducted research on the identity of English teachers through structuralism [2]. RDS Lima and his team studied the professional identity of the teaching profession constructed by different teachers in their teaching activities [3]. Teng studied the identity construction of preservice teachers, and the purpose was to study the influence of teachers' emotional changes in educational activities on teaching effect [4]. The results of Jim's research showed that the strength of expectations for teachers determined identity construction [5]. The research on teachers' identity construction has improved teachers' own cognitive structure. At the same time, the research content also includes research on the identity of native language teachers and nonnative language teachers. This provides a sufficient theoretical basis for the construction of different types of teacher identities. However, due to the lack of data and scientific method support, the practical application of research on identity construction will make its application more mere formality. It is necessary to introduce scientific research methods.

Aiming at the problem that the mentioned research on identity construction is easy to become a formality, this paper will use the feature selection algorithm to carry out the corresponding research. There have been a lot of research results on the application of this algorithm. Deniz and his team studied the application of genetic algorithms and machine algorithms to feature selection to solve some problems of binary classification [6]. Ramesh used the method of clustering to create features on the data and used different methods for the establishment of clusters to study this problem [7]. Arifused a feature selection algorithm to build a prediction system for student achievement [8]. Kalita and his team introduced a feature selection algorithm for intelligent water droplets and used a two-way limit value technique to determine the set of feature samples required for their experiments [9]. Mohammad and Alsmadi constructed a data mining detection system through feature selection, the purpose of which was to improve the accuracy of the data set [10]. These studies include many machine learning applications, some of which are used to solve classification problems, and some are used to build prediction systems. They are basically studies with many data samples. They can continuously improve the operation of feature selection algorithms, but lack of research on more theoretical research objects. It makes the fields mainly based on theory lack the support of scientific algorithms. In this paper, the feature selection algorithm is applied to the identity construction of teachers, which enables machine learning algorithms to achieve a full range of applications.

This paper adopts the feature selection algorithm in the research of teacher identity construction. The purpose of this project is to conduct more scientific research on the research object of identity construction, which is biased toward the theoretical nature, so as to better improve the teacher identity construction. In this paper, the feature selection ReliefF algorithm and Spark algorithm are used to achieve the research goal. The experiment and result analysis of teacher identity construction based on improved ReliefF algorithm in this paper. As a result, when the feature proportion of the sample is in the range of 30% and 40%, the value of AP is higher, indicating that the improved ReliefF algorithm has higher performance for the selection of samples with multiple features. The innovation of this paper is that (1) ReliefF algorithm and Spark algorithm are used in the feature selection algorithm for the identity construction of teachers, so that the research on identity construction is no longer only at the theoretical level. (2) Realizing the correlation between machine learning and the theoretical research of the research object provides more scientific algorithm support for the theoretical research object.

3. Teacher Identity Construction Method Based on Feature Selection Algorithm

3.1. The Construction of Teacher Identity. The research on teacher identity has shown great necessity for the innovative development of education, because this research includes teachers' confidence in education and teaching and their corresponding characteristics of teaching behavior. At the same time, the construction of teacher's identity is also the inner driving force that helps to influence the profession of teachers, and provides more sufficient starting power for teachers' education and teaching to realize the deepening of teachers' beliefs. The research on teacher identity construction has changed from focusing on teaching skills and professional knowledge of teachers in the past to focusing on the identity of individual teachers and is constantly exploring the interior of this profession.

3.1.1. Teacher Identity Construction. The first thing that needs to be done for the construction of teacher's identity is the identity construction of teacher's identity, and one point is developed from the nature of the profession that teachers are engaged in [11]. And it is highly consistent with the individual teachers' self-identification, but they are different from each other. The connotation of the identity construction of teacher identity is very extensive, and it is very important to the individual teacher, and the content of the identity construction of teacher identity contains several parts, and its structure can be illustrated in Figure 1:

The content in Figure 1 is often used to distinguish the object of the teaching profession, which reflects the needs of the individual teacher's life. The construction of teachers' identity is often based on the professionalism of teachers. At the same time, it will be combined with the identity of the teacher's own individual identity, and the combination of the two will eventually form a highly condensed understanding of the teacher's own nature [12]. Constructed content is based on the complete development of individual teachers, which provides a better perspective and intrinsic motivation for profound changes in education and teacher education. Teacher's identity refers to the teacher's personality, breaking the "standard" identity in the gaze and imagination of others. The process of one's own teacher identity is confirmed from the cognition and reflection of "positioning as a teacher" and "what type of teacher to become" in one's own experience.

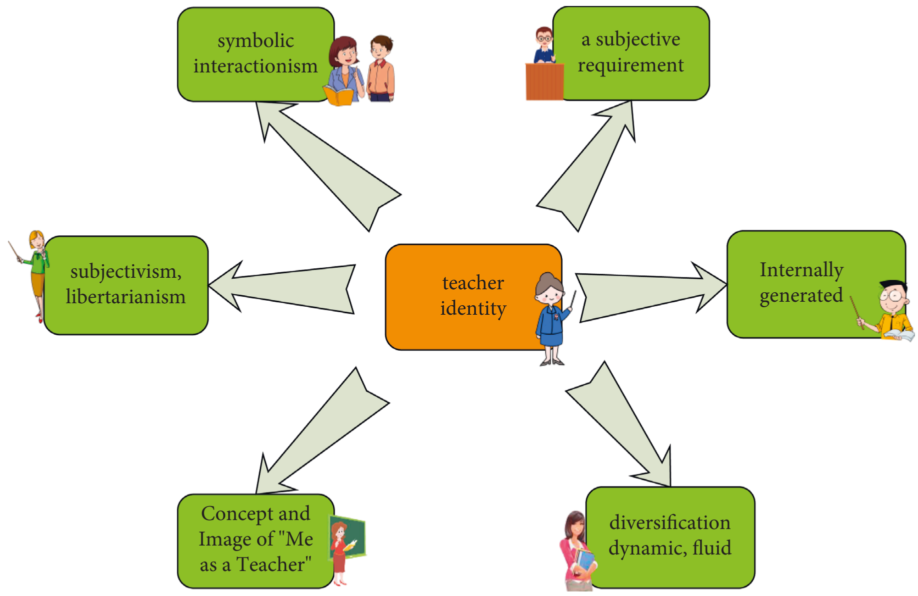

Figure 1: Construction of teacher identity content.

3.1.2. Teachers' Professional Identity Construction and Its Influencing Factors. In its meaning, the construction of teacher identity includes the individual teacher's awareness of the teacher himself and the construction of the teacher's professional identity, and the two cannot be regarded as completely independent. As this kind of understanding of the former is based on deep thinking about the individual self, its content includes the construction of teachers' self-perception of their own specialties, abilities, knowledge, and values. The entry point for the identity construction of individual teachers will start from external attention to a greater extent [13] and continue to deepen from the outside to the inside, which can be represented by Figure 2:

What is shown in Figure 2 is the identification of teachers' professional identity, which first includes the external identification of teachers' individual identity from the outside to the inside. This part belongs to social identity, and this part is the professionalism of teachers themselves that can be seen by the public for teachers themselves and for the teacher group. The teacher's inner self-construction is at the same level as the external individual's construction of the teacher himself. It has the experience of achieving professional control and being in the group through the perspective of the self and between the individual teachers and the group of teachers. It can be seen from this that the construction of teachers' professional identity has great uncertainty, and the content it contains is also diverse. Teachers' professional identity construction includes teachers' self-identity, professional identity, role identity, professional role identity, and other aspects. For the construction of the professional identity of teachers, the unique professionalism of teachers in teaching practice is involved.

The factors that have an impact on the construction of teachers' professional identity can be explored from the two levels of schools and teachers themselves. The composition of its specific influencing factors is shown in Figure 3:

The influencing factors in Figure 3 are mainly from two aspects: the teaching environment where the individual teacher is located and the individual teacher. The first factor in the former is the teaching environment where the teacher is located, which is also the classroom teaching environment. Among these factors, the teacher's control over the entire teaching classroom is the most important, which can have a long-term impact on the teacher. In addition to the abovementioned point, the influence of the teaching environment is the influence of the school environment on individual teachers, which includes the mutual influence between teachers and colleagues and student groups in the school. And the student group is very decisive for the construction of the teacher's professional identity, because the student group can show the teacher's teaching achievements well to establish the teacher's professional identity. The influence of the school environment also includes the influence of the teacher's discipline on the teacher himself. This factor is the main component of the teacher's professional identity and an entry point for the outside world to judge teachers [14]. Another major influencing factor is the teacher himself, which includes some characteristics of the teacher himself, and the influence of the events experienced by the teacher on the teacher's teaching.

3.1.3. Introduction of Different Perspectives of Teacher Identity Construction. What this paper studies is the identity construction of foreign language teachers in colleges and universities, and the former part is the clarification of the connotation of teachers' identity construction. The specific angle for the construction of teacher identity is different, because the characteristics of foreign language teaching are based on language. Therefore, the first aspect of identity construction is the professional skills of language, and the second is the knowledge of language majors, which also

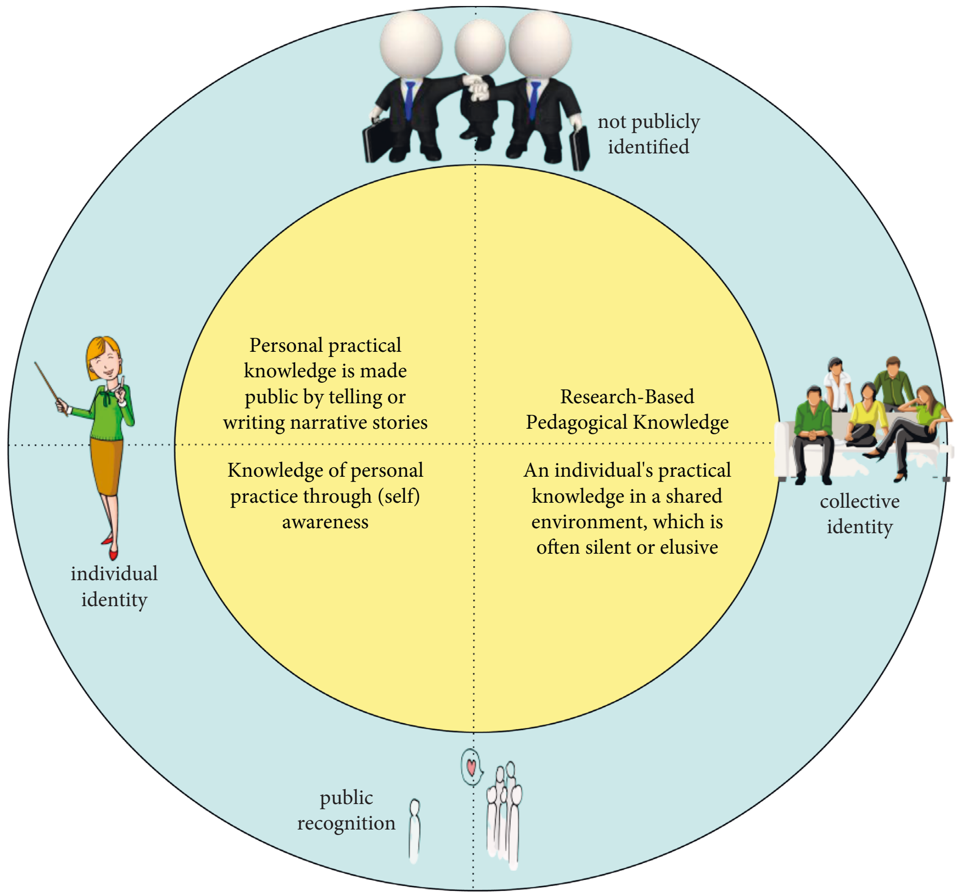

FIGURE 2: The construction connotation of teachers' professional identity.

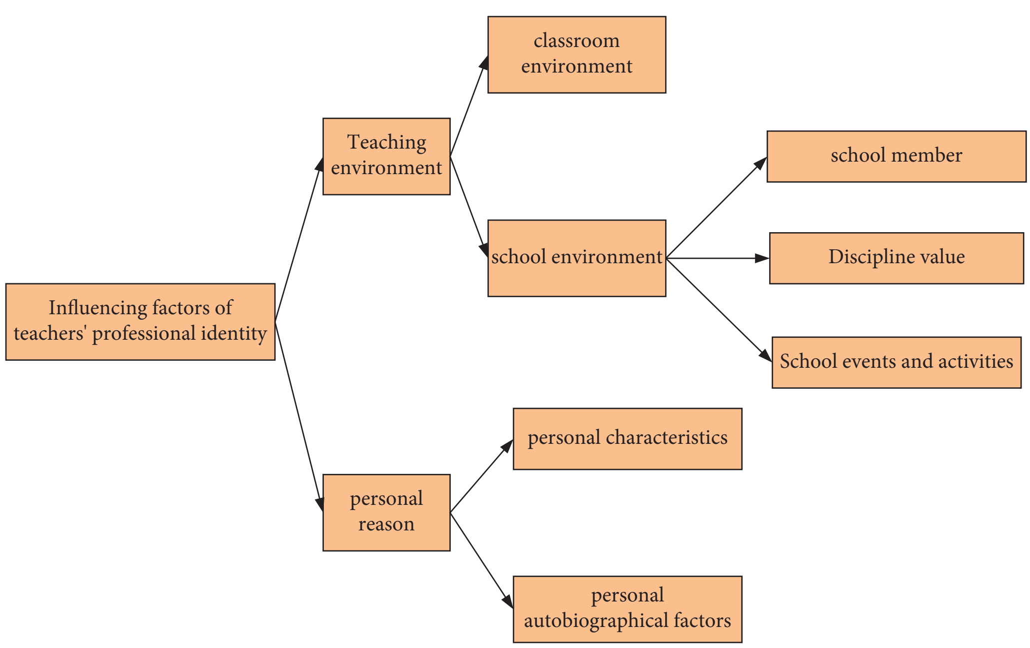

FIGURE 3: Influencing factors of teachers' professional identity construction.

contain the language views of different teachers and the practicability of foreign language teaching. Then there is the emotional tendency of individual teachers in the process of education and teaching, which is of great significance to the construction of teachers' identity. Because only the emotional orientation of teachers and their students are consistent, the results of teaching will be well manifested. For this point of view, a corresponding questionnaire is used to compare the three teachers. The results of the survey are shown in Table 1:

It can be seen from Table 1 that teacher C's emotional orientation in teaching is very consistent with that of students. The office uses five different questions to design the questionnaire. The general content of the questions is that teachers will provide effective learning plans to all students. In the process of classroom teaching, teachers will use different styles of classroom activities to enhance students' interest in classroom learning. Teachers will encourage students who encounter learning difficulties in a timely manner. Teachers' arrangements for classroom content are diverse, as well as teachers' introduction to foreign cultures. Among the three teachers involved in the questionnaire, Teacher B has the shortest teaching age, so he has the least emotional training for students, and Teacher C has participated in teaching for 15 years, so he has the most profound emotional concern for students [15]. The mean value also shows that the three teachers of different teaching ages attach great importance to the grasp of students' emotional orientation. In addition, the learning methods and strategies adopted by individual teachers play a role in the construction of teacher identity, which is very important for individual teachers' learning and growth, and it is also very important to the group of students he leads. For the different learning methods adopted, teachers will mainly affect the students' learning methods of several foreign languages. Its specific ways include cognitive strategy, regulation strategy, and resource management strategy. A survey is conducted on the learning strategies that teachers implement on students. It is mainly divided into two parts: one is students' cognitive way of foreign language learning, and the other is students' regulation way of foreign language learning. The results of the survey are shown in Figure 4.

The four factors of cognitive style in the findings in Figure 4 can be summarized as: The teacher establishes the connection in learning for the students through the common viewpoint of connection, and the teacher guides the students to summarize and think in the language learning. Teachers guide the learning methods used by students, point out the key points of the learning content, and guide students to make reasonable guesses during learning [8]. Next is the teachers' cultivation of students' foreign language culture. This perspective is also very important for the construction of teachers' identity, because as a foreign language teacher, only with a relatively high corresponding cultural literacy can he convey it to his students. A questionnaire survey is also used to examine the influence of the three teachers A, B, and C on their students' foreign language culture. The results are shown in Table 2:

Table 1: The influence of emotional orientation on teacher identity construction.

Teacher project	Q1	Q2	Q3	Q4	Q5	mean
A teacher	3	4	3	5	2	3.4
B teacher	2	3	4	4	3	3.2
C teacher	3	4	5	5	2	4.8

Table 2 shows the impact of cultural training on the identity construction of teachers. It can be seen from the table that teacher C pays the most attention to the training of students' culture, which may be related to teacher C's learning and teaching experience.

3.2. Improved ReliefF Feature Selection Algorithm. Feature selection is the product of the mature development of modern computer science and technology, because with the rapid development of computers, corresponding to the generation of many data samples. Feature selection is to summarize the set of small data samples with certain characteristics contained in the large data samples obtained by the computer. The effect of this algorithm can speed up the process of machine learnings. At the same time, this class of algorithms removes features that are irrelevant and redundant for classification. It reduces data dimensionality, avoids dimensional disasters, and can speed up the operation efficiency of learning algorithms to improve the maximum efficiency of the algorithm. This paper constructs the identity of the teacher. The content of the identity construction has been explained in the above method, which involves many characteristics of identity construction. The ReliefF feature selection algorithm here will process the extracted features. This algorithms can well select representative identity features, so as to better perform identity construction. The detailed algorithm is introduced as follows.

3.2.1. Feature Selection Principle and Classification. Feature selection is to obtain a small set of feature data samples relative to the data samples, and the application of this principle is mainly aimed at problems with classification properties. It roughly includes three steps: namely the generation of feature sample sets, evaluation, and verification of the performance of feature selection. The specific process is shown in Figure 5:

Figure 5 shows the specific operation steps of feature selection. It can be seen from the process that the normal operation of feature selection is more dependent on the method of obtaining characteristic sample data and the criteria for making certain judgments on this characteristic sample set. The method derived for the former sample set is mainly a search method of the applied data. After obtaining the characteristic data sample, it needs to make a certain judgment, which contains two different algorithm modes [16], and its corresponding structure diagram is shown in Figure 6:

Figure 6 contains two different feature selection methods. The schematic structure in Figure 6(a) is to read the features of the data samples first, and then pass the

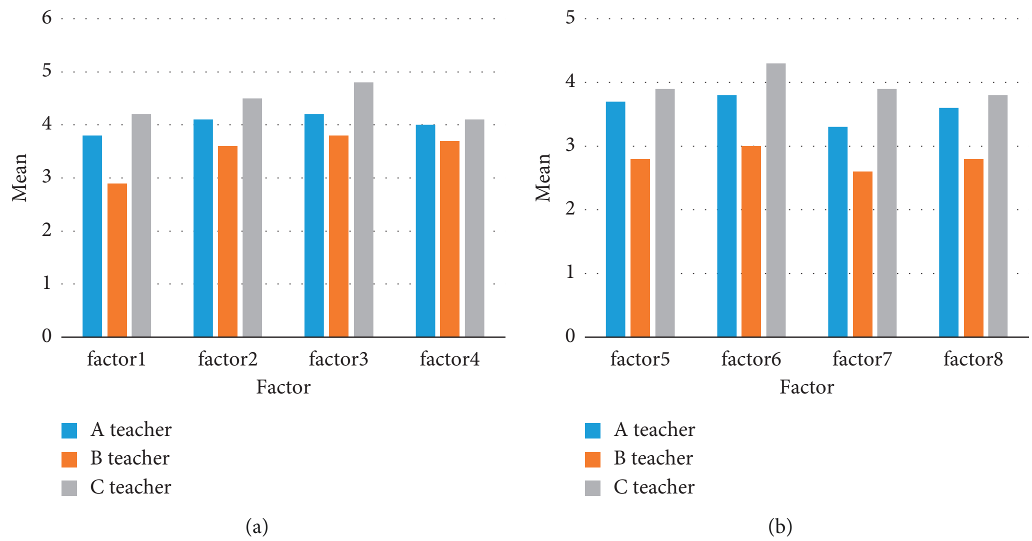

FIGURE 4: Teachers' survey on students' cognitive style and regulation style. (a) Investigation on the cultivation of cognitive style. (b) Investigation on the cultivation of regulatory style.

TABLE 2: The influence of cultural cultivation on teacher identity construction.

Teacher project	Q1	Q2	Q3	Q4	Q5	Q6	Mean
A teacher	4	5	3	5	4	3	4.0
B teacher	5	4	3	4	3	2	3.5
C teacher	4	5	5	5	4	4	4.5

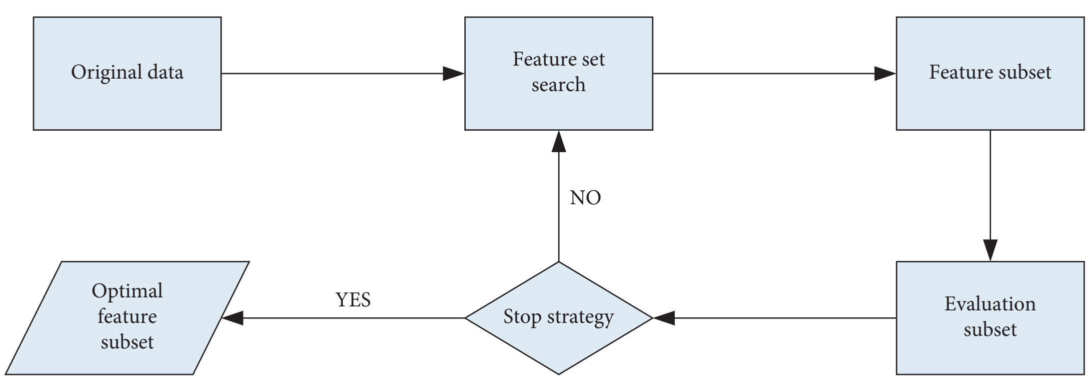

FIGURE 5: Basic flow of feature selection.

obtained feature samples to the algorithm to achieve the advance selection processing of the data. In the wrapped feature selection in Figure 6(b), the learning algorithm class is used to classify the characteristic samples so that the results will be relatively reliable.

3.2.2. Improved ReliefF Feature Selection Algorithm. The difference between the improved Relieff feature selection algorithm and the traditional algorithm is that the former can select and extract multiple features. The advantage of this is that the various characteristics of the data samples can be classified and extracted to make the operation of the algorithm more efficient. Now it is supposed that there is a data sample A, in which the sample individual is represented by a_i, and the feature number of the data sample individual q is represented by $Q_i(q)$ [17]. Assuming that the feature sample set that can be predicted by the feature selection is Q_x, then the predicted value of the q feature can be expressed by this formula:

$$Q_x(q) = \text{argmax}_{t\in\{0,1\}} H\left(M_t^q | N^q_{C_{a_x(q)}}\right). \tag{1}$$

The range of the prediction result in the formula is between 0 and 1, and the corresponding feature result is the probability of the feature in the measured data sample. $H(M_t^q | N^q_{C_{a_x(q)}})$ in the formula can only be obtained through certain mathematical changes, and the conversion formula is as follows:

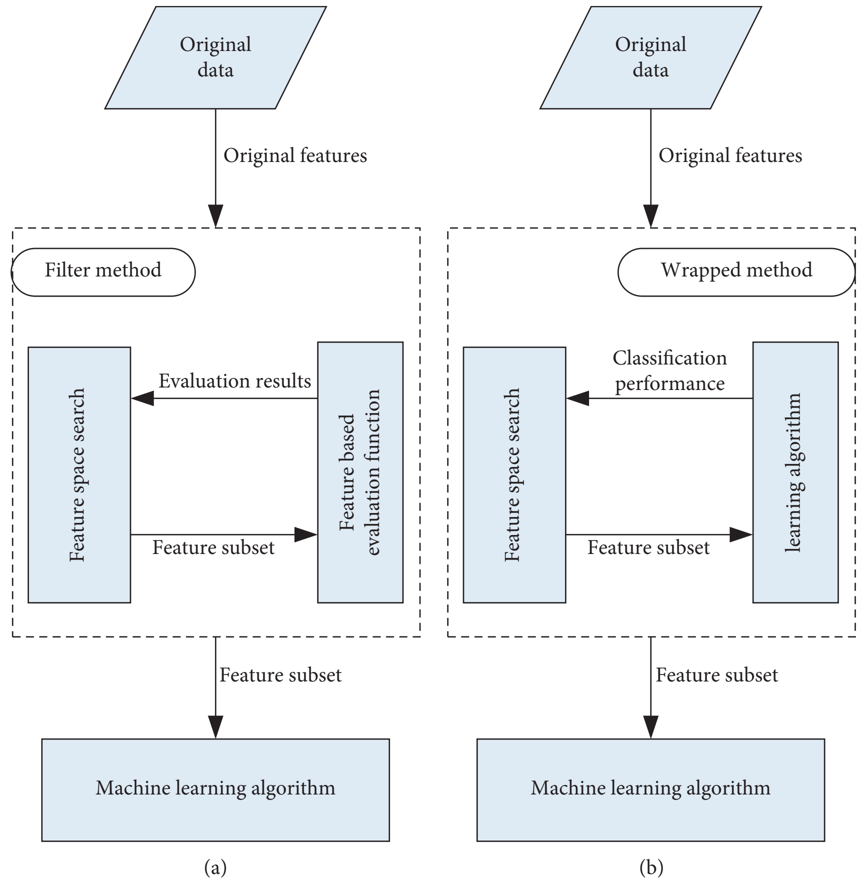

Figure 6: Two different feature selection methods. (a) Through-type feature selection diagram. (b) Wrap-around feature selection diagram.

$$H\left(M_t^q|N_{C_{a_x(q)}}^q\right) = \frac{H\left(N_{C_{a_x(q)}}^q|M_t^q\right)H\left(M_t^q\right)}{H\left(N_{C_{a_x(q)}}^q\right)}. \tag{2}$$

And $H(M_t^q)$ in the formula can be solved by the data sample, and the specific calculation process is shown in the formula:

$$\begin{aligned} H\left(M_{t=1}^q\right) &= \frac{\left(b + \sum_{i=1}^k P_i(q)\right)}{(2b+k)}, \\ H\left(M_{t=0}^q\right) &= 1 - H\left(M_{t=1}^q\right). \end{aligned} \tag{3}$$

In the formula, $H(M_t^q)$ represents the probability of the occurrence of the label q, b represents the smoothing parameter, and $H(N_{C_{a_x(b)}}^b|M_t^b)$ represents the probability after verification. The final sample prediction feature function can be expressed by the formula:

$$\begin{aligned} P_x(q) &= \operatorname{argmax}_{b\in\{0,1\}} \frac{H\left(N_{C_{a_x(q)}}^q|M_t^q\right)H\left(M_t^q\right)}{H\left(N_{C_{a_x(q)}}^q\right)}, \\ P_x(q) &= \operatorname{argmax}_{b\in\{0,1\}} H\left(N_{C_{a_x(q)}}^q|M_t^q\right)H\left(M_t^q\right). \end{aligned} \tag{4}$$

What this paper establishes is an algorithm mechanism for multiple features, and the evaluation mechanism used on a single feature selection algorithm is no longer suitable. Because in the traditional single-label learning field, metrics such as precision, precision, and recall are often used to measure model performance. However, in the multilabel classification problem, an instance can belong to multiple different categories, so it is necessary to construct a corresponding evaluation index for this problem to measure the performance of the multi-label learner. Therefore, this paper adopts another evaluation mechanism [18]. Assuming that there is an existing data sample T, a single sample in it is represented by a_i, and c is used to represent the number of features that a single sample has. The first is the evaluation of the relevant and predictive features obtained from the data samples, which can be expressed by the formula:

$$F_1(f) = \frac{1}{k}\sum_{i=1}^{k}\frac{1}{|T_i|}\sum_{r\in T_i}\frac{\left|\left\{h|\operatorname{rank}_f(a_i,h) \le \operatorname{rank}_f(a_i,l), \quad h\in T_i\right\}\right|}{\operatorname{rank}_f(a_i,l)}. \tag{5}$$

In the formula, T_i represents the set of related features in the sample, and $\operatorname{rank}_f(\bullet)$ is the function of sorting each prediction feature. The performance of this feature selection evaluation algorithm is best when the predicted value of the formula is close to 1 [19]. The second judgment algorithm is the judgment of the loss of the prediction result, which is mainly used in the case of misclassification of a single sample feature, and its expression formula is as follows:

$$F_2(m) = \frac{1}{k}\sum_{k=1}^{k}\frac{1}{D}\sum_{l=1}^{D}\left[m(a_i) \neq Y_i^l\right]. \tag{6}$$

Y_i of the formula corresponds to a set of feature samples containing two features, D represents the specific eye sample feature, and $m_l(a_i)$ represents the output predicted feature value. The smaller the result of the operation, the better the algorithm mode of feature selection in this paper. When it is necessary to judge whether the feature at the beginning of the sample features belongs to the sample, it can be expressed by the formula:

$$F_3(f) = \frac{1}{k}\sum_{i=1}^{k}\left|\left\{t \notin Q_i | f_t^h(a_i) = \max f_{t'}^h(a_i), \quad t \in T\right\}\right|. \tag{7}$$

$f_k(q)$ in the formula represents the prediction result corresponding to the feature k of the sample individual, and this algorithm is applied to the situation where errors may occur in a class of features [20]. The smaller the value of the item, the better the performance of the algorithm. When evaluating the vertical coverage of the characteristics of the data sample, the formula can be used to evaluate the performance:

$$F_4(f) = \frac{1}{k}\sum_{i=1}^{k}\max\left[\operatorname{rank}_f(a_i, l) - 1\right], \quad l \in T_i. \tag{8}$$

The smaller the value of the operation result of the formula, the better the running state of the feature selection algorithm used in this paper. In order to determine whether there is a problem with the arrangement of sample features, the formula can be used to express:

$$F_5 = \frac{1}{k}\sum_{|T_i||\overline{T_i}|^1}\left|\left\{(h,l)|\operatorname{rank}_f(a_i,h) \geq \operatorname{rank}_f(a_i,h), \ (h,l) \in T_i \times \overline{T_i}\right\}\right|. \tag{9}$$

The formula is consistent with the performance in the algorithm (10), that is, the smaller the final value is, the better the algorithm in this paper is. For the calculation formula of the ReliefF feature selection algorithm, it can be assumed that there is a data set G, $a_i \in T^h$ is the capacity of the feature of the sample, and h indicates that there are h features in the sample set. Its specific expression is as follows:

$$W(O) = \sum_{y} -\operatorname{diff}\left(O, T_i, K_y\right) + \operatorname{diff}\left(O, T_i, L_y\right). \tag{10}$$

The formula is the calculation process of the weighted value of the sample feature O, K represents the sample set similar to the sample feature, and L represents the different sample size. diff (a_1, a_2, a_3) in the formula represents the difference between a_2 and a_3 relative to feature a_1. When the distance between the sample and similar samples is smaller, the performance of the feature selection algorithm in this paper is the best.

3.3. Feature Selection Based on Spark Algorithm. The application of this algorithm is for the further improvement of the feature selection algorithm of ReliefF in 2.2, and the improvement method needs to refer to the principle of another ReliefF algorithm. Spark algorithm is used to solve multiway processing objects. After the above feature selection algorithm extracts and optimizes the feature elements of teacher identity, what is needed is to perform more comprehensive algorithm processing on the characteristics of the selected teacher identity. In this way, a more precise construction of the teacher's identity is carried out. For the referenced algorithm object, it is also transformed according to the process of the relief algorithm [21], and its main process is shown in Figure 7:

The flow of the reference ReliefF algorithm in Figure 7 is partly similar to the original ReliefF algorithm. For the calculation of the distance between samples, the formula can be used:

$$f_d(a) = \frac{1}{2}\left(\left\|a - a^x\right\| - \left\|a - a^y\right\|\right). \tag{11}$$

a^x and a^y in the formula, respectively, represent the nearest feature samples with similarity and the nearest feature samples with dissimilarity in the detected data samples. Suppose that the set with r sample individuals is represented by G^r, and the original sample set is mapped to form a set with a certain space capacity. The formula is as follows:

$$a_i' = |a_i - a_i^x| - |a_i - a_i^y|. \tag{12}$$

The a_i' in the formula represents the i th feature of the sample a' in the sample feature capacity set G'^r formed by the mapping. a_i^x and a_i^y are the same as the i-th feature corresponding to a^x and a^y in the original sample. From this, the sample generated by the mapping has all the characteristic elements of the original sample. The calculation formula can only be performed for one adjacent feature. In order to reduce the generation of errors and the influence of noisy data or abnormal data on space conversion, the formula can be used for calculation:

$$a_i' = \sum_{j=1}^{n} a_i - a_i^{xj} - \sum_{j=1}^{m} a_i - a_i^{yj}. \tag{13}$$

In the formula, n represents the number of samples with the same feature, while m represents the number of samples with different features, and the sum between the two is the total number of tested samples. a_i^{xj} represents the coordinate value of the dissimilar adjacent samples of the sample a in the sample, and a_i^{yj} represents the coordinate value of the adjacent samples with similar characteristics. Referring to the relief algorithm, the similarity between samples is distinguished by adding different weights to the samples. The solution formula for a sample weight can be expressed as follows:

$$P(a) = \frac{1/\overline{D}(a')}{\sum_{j=1}^{m} 1/\overline{D}\left(a_j'\right)}. \tag{14}$$

For $\overline{D}(a_j')$ in the formula, it can be calculated by the formula, and the formula expression is as follows:

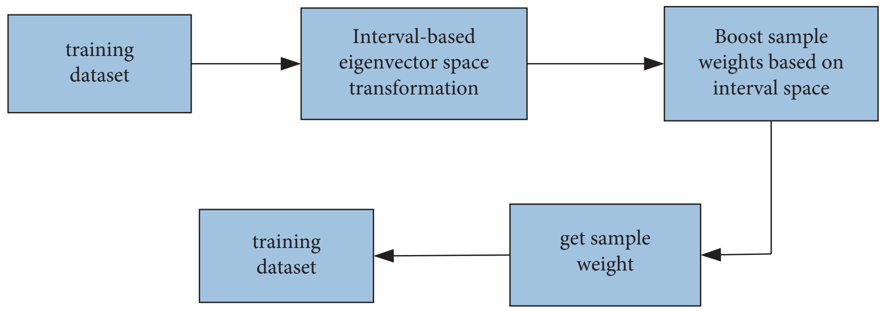

Figure 7: Referring to the operation flow of the relief algorithm.

$$\overline{D}\left(a_j'\right) = \frac{1}{m-1}\sum_{q=1,a_q'\neq a'}^{m-1}\left\|a' - a_q'\right\| \tag{15}$$

The weight frame of the reference algorithm can be obtained by superimposing the feature capacity of the sample and the weighted value of the sample. The algorithm can finally get the weight of the sample. In order to achieve the stable operation of the feature selection algorithm, it is combined with the ReliefF algorithm. The most primitive formula of the ReliefF algorithm is as follows:

$$Q\left(A_i\right) = \frac{1}{\mathrm{mk}}\sum_{u=1}^{m}\sum_{j=1}^{k}\left(\left|a_{u,i} - a_{u,i}^{xj}\right| - \left|a_{u,i} - a_{u,i}^{yj}\right|\right). \tag{16}$$

In the formula, $a_{u,i}$ indicates that the coordinate value corresponding to the sample is i, $a_{u,i}^{xj}$ indicates that the dissimilar sample number is the adjacent sample coordinate value of j, and m in the formula indicates the number of samples. Finally, the stable feature selection formula can be obtained:

$$Q\left(A_i\right) = \sum_{u=1}^{m} q_j \sum_{j=1}^{k}\left(q_u^{xj}\left|a_{u,i} - a_{u,i}^{xj}\right| - q_u^{yj}\left|a_{u,i} - a_{u,i}^{yj}\right|\right). \tag{17}$$

In the formula, q_j represents the weight of the sample, q_u^{xj} represents the weight of the sample with the sample number j whose samples are dissimilar, and q_u^{yj} represents the weight of the sample with similarity. It can be seen from the formula that the size of the weights and the degree of correlation of the sample features are positively correlated [22].

4. Experiments on Teacher Identity Construction with Feature Selection Algorithms

4.1. Investigation and Results of Teacher Identity Construction. The survey on the construction of teachers' identity is mainly in the form of questionnaires, taking foreign language teachers in colleges and universities as the research object of this paper. The purpose is to explore the main problems of the identity construction of foreign language teachers in colleges and universities, so as to promote the foreign language learning of college students. Through the survey, some information about the teachers in Table 3 can be obtained.

The total number of teachers in Table 3 is 200. It can be seen from the table that most of the young foreign language teachers are younger teachers with shorter teaching time. This is also in line with the characteristics of the teaching profession, and in foreign language teaching, teachers who teach English account for a relatively large number of teachers. For these teachers, the questionnaire uses five aspects to investigate the identity construction of teachers, including teachers' self-image, self-evaluation, professional status, professional motivation, professional emotion, and professional expectation. The evaluation of young teachers of men and women in various aspects is different, and the results are shown in Table 4:

From Table 4, the difference between teachers' identity construction and male and female teachers is very small. Only in the teacher's identity construction, there is a big difference in the teacher's own image and major. The average rating of young female teachers for teachers' self-image is 3.78, which is 0.53 higher than that of young male teachers, and the gap between female teachers and male teachers is also 0.35 in the evaluation of teachers' professional prospects. This shows that the identity construction standards of young women foreign language college teachers are mostly based on their majors. In addition, the influence of teachers' age and teaching years on the construction of teacher identity is also investigated, and the results are shown in Tables 5 and 6:

Tables 5 and 6 are one of the influencing factors for the construction of teacher identity. In Table 5, the influence of teacher age on the construction of teacher identity is small, but there are certain differences between different age groups. Teachers aged 31–40 have a greater role in the construction of their own professional identity. It can be seen from Table 6 that the teaching years of teachers have no obvious influence on the identity construction of teachers. Then, a data analysis of two factors is conducted on the influence of teacher self-evaluation on teacher identity construction. One factor is the knowledge that teachers themselves think is currently lacking and the factors that hinder teachers' growth. The results are shown in Figure 8:

From Figure 8(a), teachers of different ages who lack different knowledge have different needs. Among them, the number of teachers aged 20–25 who think they lack relevant knowledge of teaching methods and professional knowledge is the largest, reaching 61 and 22, respectively. Figure 8(b)

Table 3: Basic information of young foreign language teachers in colleges and universities.

Gender	Classification	Number of people (person)	Percentage (%)
Age	Male	36	21.95
Teaching age	Female	164	78.05
Education	20–25 years old	108	54
Teaching subjects	26–30 years old	48	26
Gender	31–35 years old	36	20
Age	35–40 years old	8	4
Teaching age	3–5 years	85	42.5
Education	6–10 years	63	31.5
Teaching subjects	11–15 years	31	15.5
Gender	16–20 years	21	10.5
Age	Undergraduate	82	41
Teaching age	Master degree and above	118	59
	English	152	76
Education	French	24	12
	Russian	24	12

Table 4: Survey results of the identity construction of young male and female teachers.

Project	Gender	Mean	SD
Self-image	Male	3.25	0.52
	Female	3.78	0.35
Self-evaluation	Male	3.41	0.57
	Female	3.56	0.61
Professional status	Male	3.72	0.54
	Female	3.74	0.72
Professional emotion	Male	3.18	0.71
	Female	3.53	0.95
Professional expectations	Male	3.11	0.61
	Female	3.27	0.73

Table 5: Analysis of variance of teacher age on teacher identity construction.

Project	Mean	SD
Self-image	3.67	0.51
Self-evaluation	3.87	0.53
Professional status	3.69	0.49
Professional emotion	3.21	0.37
Professional expectations	3.05	0.61

Table 6: Variance analysis of teaching years on teacher identity construction.

Project	Mean	SD
Self-image	3.72	0.53
Self-evaluation	3.95	0.57
Professional status	3.73	0.53
Professional emotion	3.26	0.41
Professional expectations	3.11	0.63

shows that teachers' income, social status, and work pressure are the most important factors in the group of teachers aged 20–30 among the growth-impeding factors for the construction of teachers' identity. In this age group, work pressure has an impact on the construction of teachers' identity, reaching 95 people, which is close to the general number of survey respondents. This also shows that reasonable adjustment of teachers' work pressure is conducive to the construction of their identities.

4.1.1. Experiments and Results of Teacher Identity Construction Based on Improved ReliefF Algorithm. In order to verify the improved relief algorithm, it is necessary to perform classification and selection operations on various features for the relevant sample sets. In addition to using the improved relief algorithm, this paper also selects two other algorithms for comparison. The selected feature sample sets are A and B. Finally, the experimental results can be obtained as shown in Figure 9:

It can be seen from Figure 9 that for different selection algorithms adopted for different sample sets, the verification results have different changes. Among them, the distribution method of the number and the improved relief algorithm have a smaller feedback range for the data samples. However, the weight algorithm in the unit has a larger variation range for the verification result of the sample, which shows that the operation of the weight algorithm in the unit is not stable enough. The AP of the improved relief algorithm is better than the number allocation algorithm most of the time, and when the feature ratio of the sample is in the range of 30% and 40%, the value of the AP is higher. It can be shown that the improved relief algorithm has higher performance for the selection of samples with multiple features.

4.1.2. Experiments and Results of Teacher Identity Construction Based on Spark Algorithm. The Spark algorithm is obtained by optimizing the improved relief algorithm. In the process of operation, it is mainly to measure the running time of the algorithm. In the experiment, the aggregation degree of the algorithm is set to 25, and when the memory of the execution components of the algorithm is 35 G, the number of execution components will have a greater impact

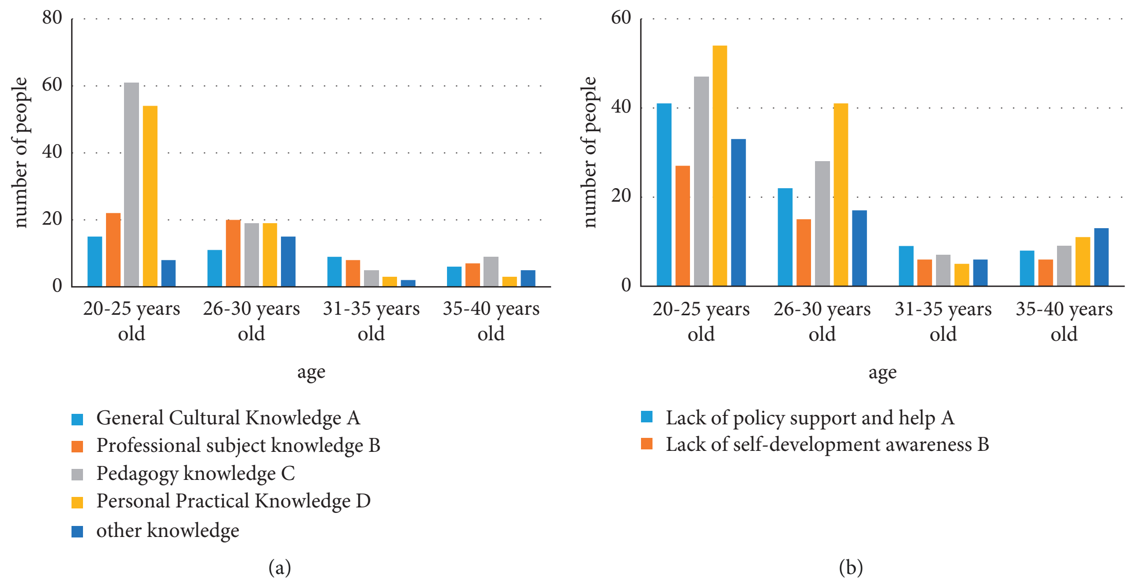

Figure 8: Data analysis of teacher self-evaluation for identity construction. (a) Data analysis of lack of knowledge in self-evaluation of teacher identity construction. (b) Data analysis of growth-impeding factors in the self-evaluation of teacher identity construction.

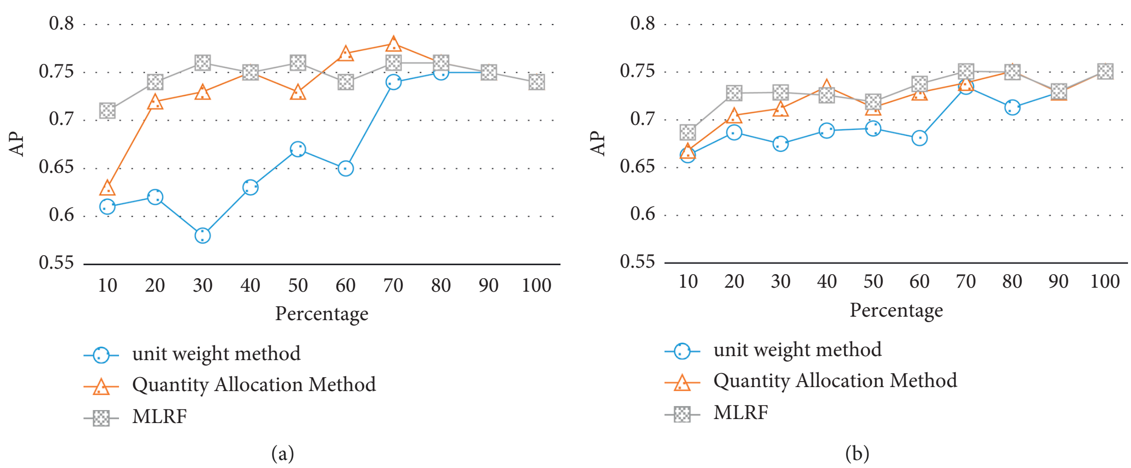

Figure 9: Test results of different feature selection algorithms on different sample sets. (a) Verification results of different methods on set A. (b) Verification results of different methods on set B.

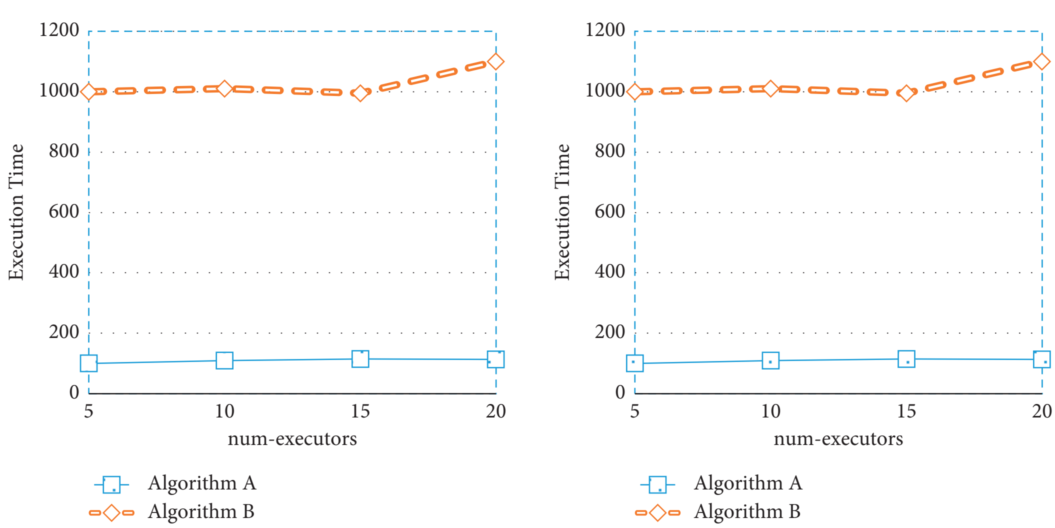

Figure 10: The effect of the number of execution parts on the running time of the algorithm.

on the operation of the algorithm, and the corresponding results are shown in Figure 10:

It can be seen from Figure 10 that there is a large gap between the running time of Algorithm A and Algorithm B, and the time used by the latter is about 10 times longer than that of the former. This is the experimental result in the case of specific parameters, except that the number of execution components will have an impact on the running time of the algorithm, the impact of changes in other parameters is very small. In this regard, it is necessary to debug the parameters of the Spark algorithm, and setting the aggregation degree of the sample to 25 will get a better running effect.

5. Conclusion

This paper studied the identity construction of foreign language teachers in young colleges and universities. The main purpose of this research was to provide a better foundation for the development of teachers. In traditional classroom education, people often only focus on teachers' professional knowledge and teaching methods, and lack of understanding of the connotation of teachers' identity construction. This is of great significance for changing the previous teaching mode. This paper used two different feature selection algorithms to study the construction of teacher identity and extracted and analyzed the factors that have an important impact on the identity construction of teacher groups, and different algorithms had different effects on the processing of the sample data obtained from the survey. Through the introduction of such algorithms as feature selection, the research on teacher identity construction can be made complete and more valuable to realize the self-improvement of the teacher group.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This research study was sponsored by The 2021 Youth Academic Project of Guangzhou Huashang College. The name of the project is Multi-approach Study on the Self-development Ability of Young Foreign Language Teachers from the Perspective of New Liberal Arts (project no. 2021HSQX43). Grateful for the support to this article.

References

[1] C. Zhang and Y. Zhang, "Language teacher identity construction: insights from non-native Chinese-speaking teachers in a Danish higher educational context," *Global Chinese*, vol. 4, no. 2, pp. 271–291, 2018.

[2] N. T. Zacharias, "The teacher identity construction of 12 asian NNES teachers in TESOL graduate programs," *Journal of Asia Tefl*, vol. 7, no. 2, pp. 177–197, 2018.

[3] R. D. S. Lima, S. P. Pereira, and R. Kanaane, "The teacher identity construction in the context of secondary education integrated to professional and technological education," *Boletim Academia Paulista de Psicologia*, vol. 37, no. 93, pp. 322–352, 2017.

[4] M. F. Teng, "Emotional development and construction of teacher identity: narrative interactions about the pre-service teachers' practicum experiences," *Australian Journal of Teacher Education*, vol. 42, no. 11, pp. 117–134, 2017.

[5] M. K. Jim, "Identity construction in learning English academic writing in a Japanese university," *The Journal of AsiaTEFL*, vol. 14, no. 2, pp. 228–243, 2017.

[6] A. Deniz, K. He, T. Dokeroglu, and A. Cosar, "Robust multiobjective evolutionary feature subset selection algorithm for binary classification using machine learning techniques," *Neurocomputing*, vol. 241, pp. 128–146, 2017.

[7] E. A. C. R. Ramesh, "Cluster optimization for boundary points using distributive progressive feature selection algorithm," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 2, pp. 3320–3328, 2021.

[8] A. Arif, A. Jahan, M. I. Mau, and R. Tummarzia, "An improved prediction system of students' performance using classification model and feature selection algorithm," *International Journal of Advances in Soft Computing and Its Applications*, vol. 12, no. 1, pp. 162–177, 2021.

[9] D. J. Kalita, V. P. Singh, and V. Kumar, "Two-way threshold-based intelligent water drops feature selection algorithm for accurate detection of breast cancer," *Soft Computing*, vol. 26, no. 5, pp. 2277–2305, 2021.

[10] R. M. A. Mohammad and M. K. Alsmadi, "Intrusion detection using Highest Wins feature selection algorithm," *Neural Computing & Applications*, vol. 33, no. 2, pp. 1–12, 2021.

[11] X. Chen and C. X. Zhang, "Identity construction through tourism education: voice from international students," *Critical Tourism Studies Proceedings*, vol. 2019, no. 1, p. 9, 2019.

[12] K. Jeongyeon and S. H. Young, "Beyond nativeness versus nonnativeness in the construction of teacher identity in the context of Korean as a foreign language," *The Sociolinguistic Journal of Korea*, vol. 27, no. 2, pp. 25–50, 2019.

[13] E. Boranda, A. Zft, and Y. Kaygusuz, "Development of majority vote ensemble feature selection algorithm augmented with rank allocation to enhance Turkish text categorization," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 29, no. 2, pp. 514–530, 2021.

[14] Z. abbasi, M. rahmani, and H. ghaffarian, "IFSB-ReliefF: a new instance and feature selection algorithm based on ReliefF," *Signal and Data Processing*, vol. 17, no. 4, pp. 49–66, 2021.

[15] S. S. Shreem, H. Turabieh, S. Al Azwari, and F. Baothman, "Enhanced binary genetic algorithm as a feature selection to predict student performance," *Soft Computing*, vol. 26, no. 4, pp. 1811–1823, 2022.

[16] M. D. D. Lima, J. D. O. R. E. Lima, and R. M. Barbosa, "Medical data set classification using a new feature selection algorithm combined with twin-bounded support vector machine," *Medical, & Biological Engineering & Computing*, vol. 58, no. 3, pp. 519–528, 2020.

[17] P. García-Díaz, I. Sánchez-Berriel, J. A. Martínez-Rojas, and A. M. Diez-Pascual, "Unsupervised feature selection algorithm for multiclass cancer classification of gene expression RNA-Seq data," *Genomics*, vol. 112, no. 2, pp. 1916–1925, 2020.

[18] R. Shang, J. Song, L. Jiao, and Y. Li, "Double feature selection algorithm based on low-rank sparse non-negative matrix factorization," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 2, pp. 1–18, 2020.

[19] M. K. Uar, "Classification performance-based feature selection algorithm for machine learning: P-score," *IRBM*, vol. 41, no. 4, pp. 229–239, 2020.

[20] O. Selvi and İ. Huseyinov, "A novel algorithm for feature selection based on geographic distance metric: a case study of streamflow forecasting of Austria's water resources," *International journal of Environmental Science and Technology*, vol. 17, no. 1, pp. 295–308, 2020.

[21] T. B. Alaku and I. Turkoglu, "Feature selection with sequential forward selection algorithm from emotion estimation based on EEG signals," *Sakarya University Journal of Science*, vol. 23, no. 6, pp. 1096–1105, 2019.

[22] M. Hashem, A. Bashandy, and S. Shaheen, "Improving anomaly detection in BGP time-series data by new guide features and moderated feature selection algorithm," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 27, no. 1, pp. 392–406, 2019.