Research on Pattern Recognition Method of Online English Education Based on Feature Self Learning

Xiao-xiao Duan[1](✉) and Ping Duan[2]

[1] International School, Chongqing Vocational Institute of Engineering, Jiangjin 402260, China
[2] Big Data and Internet of Things School, Chongqing Vocational Institute of Engineering, Jiangjin 402260, China

Abstract. Aiming at the problem of low accuracy of pattern recognition in online English education, a pattern recognition method based on feature self-learning is proposed. Using feature self-learning algorithm and convolution neural network, the feature extraction model of online English education is established, and the basic features and depth features of online English education are extracted. Z transform is introduced to standardize the feature data set of English online education, and Laplace gradient function is used to clean the wrong and missing data of English online education features. The fuzzy logic theory is used to infer the important parameters of English online education. According to the determined characteristics and parameter values of English online education, the difference between the characteristics and parameter values of English online education is calculated, and the mode of English online education is identified. The experimental results show that: This study of English online education pattern recognition method, can identify all the patterns of English online education, and English online education pattern recognition accuracy is high.

Keywords: Feature self-learning · English online education · Pattern recognition

1 Introduction

The deep integration of education and information technology has been the focus of our country for a long time, and the government has also issued many relevant policies to support online education. In this context, with the rapid development of online education, online learning has gradually become a common way of learning. Compared with traditional education, online education has the advantages of unlimited time, geographical location and resource sharing [1]. However, due to the late start of online education in China, there is a certain gap compared with developed countries. At the same time, there are also some problems such as the lack of monitoring and management of the platform, the same education mode, and many networks only stay at the level of resource sharing. Pattern recognition method, as a method to determine whether the element (the discriminated object) belongs to a certain set (a certain state), has been widely used

W. Fu et al. (Eds.): eLEOT 2021, LNICST 389, pp. 266–277, 2021.
https://doi.org/10.1007/978-3-030-84383-0_23

in various fields [2]. For pattern recognition methods at home and abroad, k-nearest distance discrimination method, various clustering analysis methods, artificial neural network and support vector machine methods have been developed [3–5]. However, the pattern recognition method of the above research does not consider the relationship between the classification features. In fact, there is a certain relationship between the extracted classification features. Therefore, the above research is not perfect, more or less has certain defects, which seriously restricts its further application in various fields. At present, many disciplines have adopted online education mode for teaching, among which English online education mode has the highest similarity, and it is difficult to judge which education mode is the most effective. Therefore, this paper introduces the feature self-learning algorithm, makes use of its excellent feature extraction ability for deep subspace and the advantages of running time to study the pattern recognition method of English online education, and puts forward the research topic of pattern recognition method of English online education based on feature self-learning.

2 Research on Pattern Recognition Method of English Online Education Based on Feature Self Learning

2.1 Establish Feature Extraction Model Based on Feature Self-learning

Based on the data collected by the mobile terminal, the basic features of English online education are obtained by simple calculation. Convolution neural network further processes the features of basic English online education and extracts the depth features of English online education. Based on the feature self-learning algorithm, the feature extraction model is established. The whole framework is generally composed of four parts as Fig. 1.

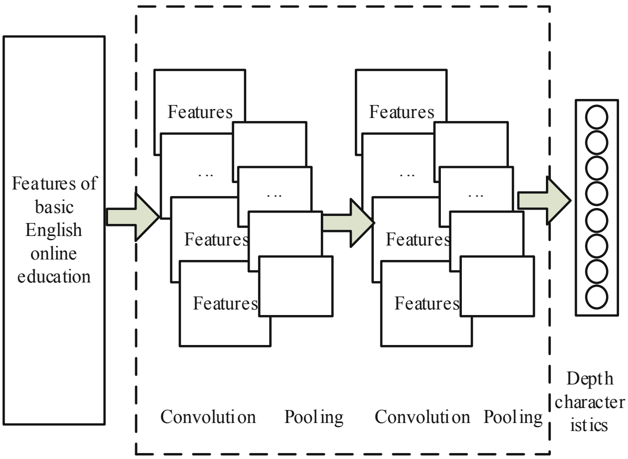

Fig. 1. Feature extraction model

In the feature extraction model of English online education shown in Fig. 1, the feature self-learning algorithm needs convolution in convolution neural network to get the features of English online education. In order to solve this problem, a basic artificial feature time series is composed of the sampling points of the English online education mode segment and the time sequence relationship. The convolution neural network is divided into two steps to pool the feature sequence. Firstly, the 2-D convolution network is used to map the sequence features, and the pooling layer is used to compress them. Then, the 1-D convolution is performed on the compressed sequence, and the 1-D depth features are obtained by pooling again. The convolution process is as Fig. 2.

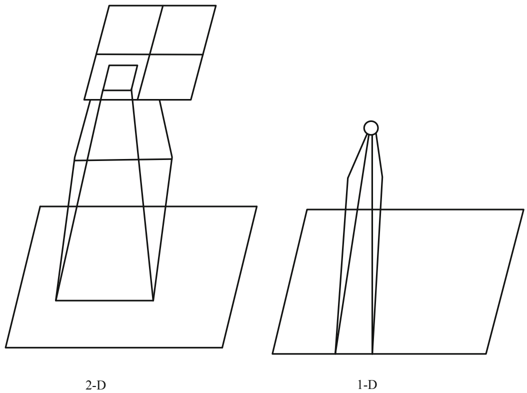

Fig. 2. 2-D and 1-D convolution processes

Pooling is another important operation in convolutional neural network processing. It can compress the feature maps generated after convolution. There are two common pooling strategies: maximum and average. Using the maximum pooling policy, the maximum value in the activation unit is used as the pooling output; using the average pooling policy, the average value of the activation unit is used as the output.

The feature self-learning algorithm calculates the features of basic English online education, and then uses convolution neural network to learn the features of basic English online education based on BHF. Finally, the algorithm returns depth features for further processing, The calculation process is as follows:

(1) Input English online education segment;
(2) Extract the features in the fragment;
(3) Set the threshold of segment feature extraction;
(4) According to the threshold, extract the features of English online education segments repeatedly;
(5) Output the basic feature extraction results of English online education segment.

Based on the above contents, the feature extraction of English online education is completed. According to the features of English online education extracted in Fig. 1, the feature database of English online education can be constructed, and the English online education mode can be identified according to the features of English online education.

2.2 Construction of English Online Education Feature Database

Based on the feature extraction model of English online education shown in Fig. 1, the extracted feature data of English online education is the original feature data set of English online education, including attributes and their corresponding attribute values. The data set contains duration data and accuracy data, both of which are continuous data. It will be inconvenient in the process of pattern recognition in English online education [6]. Therefore, it is necessary to transform and standardize the data to make all attribute values fall into the limited space, which is conducive to data mining and analysis. Therefore, Z-transform is introduced to standardize the feature data set of English online education, so that all the data fall into one range:

$$z = \frac{(x - \mu)}{\sigma} \tag{1}$$

In formula (1), z represents the converted data; x represents a certain data of the column of data to be converted; μ represents the average of this column of data; σ represents the standard deviation [7].

Since the final data type of learning style is discrete value, the two types of data need to be analyzed later, so the online learning behavior data set also needs to be discretized. Therefore, the transformed data set is equally divided into five intervals, and finally the discrete value data set of all online learning behaviors is obtained.

The standardized data obtained from the conversion of formula (1) needs to be converted and cleaned accordingly. Therefore, the data are processed as follows. For information processing dimensions, use H for active type, B for balanced type, and C for contemplative type; for information perception dimension, use G for sensory type and B for balanced type; Z for intuitive type; for information input dimension, S represents the visual type, B represents the balanced type, and Y represents the verbal type; for the information understanding dimension, X represents the sequence type, B represents the balanced type, and Z represents the comprehensive type.

At this time, Laplace gradient function is used to calculate the second derivative of the data, so as to measure the data missing and error. If the variance of Laplacian operator of English online education data is lower than a fixed value, it means that the change of English online education data is not obvious, and the English online education data can be considered as missing error data. According to the Laplacian operator, cleaning the sample data of English online education, there are errors, missing data, all stored in the database, that is to complete the construction of English online education feature database.

2.3 Infer the Important Parameters of English Online Education

Fuzzy logic theory is introduced to infer the important parameters of online English education. Membership function is defined to represent the relationship between input and output variables in the inferential process. Finally, simple and clear results are output through fuzzy inference. The reasoning process is as follows:

(1) Determine the membership function of input variables, so as to determine the rules in the rule base that the input variables can trigger. Because there are a large number of rules in the rule base, when each input variable is input, it is impossible to trigger all the rules, only some rules related to the corresponding input variables will be triggered;
(2) Judge the credibility of the rules triggered by each input variable. We already know some rules that may be triggered by the input variable. At this time, we need to calculate the credibility of each rule triggered according to the actual situation, find the most likely rule, and then carry out the next step of fuzzy reasoning to judge the possible results;
(3) To judge the final output of the fuzzy reasoning system, we need to comprehensively judge the credibility of the rules triggered by the previous input variables to calculate and determine the final fuzzy reasoning result.

Based on the above three steps, determine the membership function of the input variable. $u(x)$ will be used to represent the membership function of the input function relative to the fuzzy set. Then the value of $u(x)$ is closer to 0, indicating that the variable x is related to any feature of English online education in the domain. The relationship of the set A, the lower the degree of implication, that is, the lower the degree that x belongs to the feature set A of English online education. Conversely, when the value of $u(x)$ tends to 1, the higher the degree of implication between the x variable and the feature set A of English online education, that is, the higher the degree of membership [8, 9].

On this basis, a fuzzy logic rule library rule is determined, and its inference composition rules are as follows: suppose r represents a fuzzy relationship in uv, where u represents an element in the feature set of English online education; v represents the feature of English online education Fuzzy collection. And x represents a fuzzy subset in the domain U. Then, with the help of x inference, the fuzzy subset $y = xr$ in v is derived. This inference method is called Sup-Star inference synthesis rule, where xr is the Sup of x and r-Star synthesis. The specific synthesis rules are as Table 1.

2.4 Identify English Online Education Mode

Suppose that the matrix of all the original online English education models is $\underline{\varphi}$ and $\varphi = [\varphi_1, \varphi_2, \ldots, \varphi_c] \in R^{mN}$, where c is the number of categories of online English education models, m is the number of training samples, and N is the total number of samples. According to the hypothesis object, the model of English online education is identified. Using the above three contents, the characteristics and parameters of English online education are extracted, and the neutral approximate sample is obtained. The difference between the test sample and the neutral approximation sample is made to obtain the difference of the English online education characteristics of the test sample.

The neutral approximate sample is a linear approximate representation of the neutral feature set of English online education, and its vector is represented as y_n, then:

$$y_n = \varphi \hat{w}_q \tag{2}$$

Table 1. Rules of fuzzy logic theory rule base

GMP	Premise 1	*X* is *A'*
	Premise 2	If *X* is *A*, then *y* is *B*
	Conclusion	*Y* is *B'*
	Fuzzy inference	$B' = A'(A \rightarrow B) = A'r$
GMT	Premise 1	*Y* is *B'*
	Premise 2	If *X* is *A*, then *y* is *B*
	Conclusion	*X* is *A'*
	Fuzzy inference	$A' = (A \rightarrow B)B' = rB'$

Note: *r* represents the parameters containing fuzzy implication relations. The widely used fuzzy operations include fuzzy implication product operation and fuzzy implication minimum operation. Any form of fuzzy operation reasoning can be realized with the help of *r*; other letters do not have practical meaning, only indicate the reasoning process of fuzzy logic theory rule base

In formula (2), φ represents the neutral feature set of English online education, w_{q1} represents the weighted vector, the lower corner represents the q feature of English online education, which is obtained by the normalized least square optimization formula, then:

$$w_{q1} = \arg\min\left\{\left\|y_q - \varphi w_q\right\|_2^2 + \lambda\left\|w_q\right\|_2^2\right\} \tag{3}$$

In formula (3), λ represents the normalization parameter, y_q represents the test sample vector, $\|\cdot\|_2$ represents the norm vector, w_q represents the q feature vector of English online education, which can be obtained by the following formula:

$$w_q = \left(\varphi^T\varphi + \lambda I\right)^{-1}\varphi^T y_q \tag{4}$$

In formula (4), I represents the identity matrix, T represents the matrix conversion.

At this time, the λ in Eq. (5) is set to 0.1, and the differential test sample y can be expressed as:

$$y = \lambda\left(y_q - y_n\right) \tag{5}$$

English online education feature difference set D contains N difference training sample vectors, namely $D = [v_1, v_2, \ldots, v_c]$, which also contains C English online education feature categories. Each difference training sample vector can be expressed as v_i, where i represents differential training sample [10]. Assuming that $v_s \in R$ is a random training sample in the original English online education, the differential training sample v_i is also obtained by calculating the difference between the training sample v_s and its corresponding neutral approximate sample v_n. The solution method is the same as the differential test sample y.

Assume that the English online education mode corresponding to any difference vector v_i is I, where $I_j = (1, 2, \ldots, n)$ represents the I column of j. If I is an even number, then the English online education feature vectors s_1 and s_2 of the English online education model I can be expressed as:

$$s_1 = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_{\frac{j}{2}} \end{bmatrix}$$
$$s_2 = \begin{bmatrix} I_j \\ I_{j-1} \\ \vdots \\ I_{\frac{j}{2}+1} \end{bmatrix} \tag{6}$$

In formula (6), s_1 is obtained by concatenating the parameter values from the first column to the $j/2$ column of the English online education model I. Similarly, s_2 is obtained by concatenating the parameter values from the last column to the $j/2 + 1$ column of the English online education model I [11].

According to the difference value obtained above, the number of English online education features and parameter values is determined to be p. At this time, X is used to represent English online education features and parameter values, then $X = [X_1, X_2, \ldots, X_p]$, for any one of the English online education features and parameter values English online education uses eigenvalues to predict X_i, where $j \neq i$ is:

$$X_i = f\left(X_j, b_0, b_j, b_{jj}, b_{jk}\right) + e \tag{7}$$

In formula (7), X_i represents the predicted variable; X_j represents the predictor variable; e represents the prediction error; $b_{0,}$ $b_{j,}$ $b_{jj,}$ $b_{jk,}$ represents the English online education parameter [12].

According to formula (7), using the RBF network model with a three-layer network of input layer, hidden layer and output layer to identify the English online education mode, there are:

$$X_j = \sum_{r=1}^{h} \omega_r \cdot \phi_r(X_i) \tag{8}$$

In formula (8), r represents the Gaussian node; h represents the number of Gaussian nodes; ω_r represents the center vector of the r Gaussian node in the hidden layer of the function; φ_r represents the actual value of the r Gaussian node in the hidden layer of the function output [13]. Through the above steps, the recognition result of English online education mode can be obtained. The pattern recognition process of online English education based on feature self-learning is shown in Fig. 3.

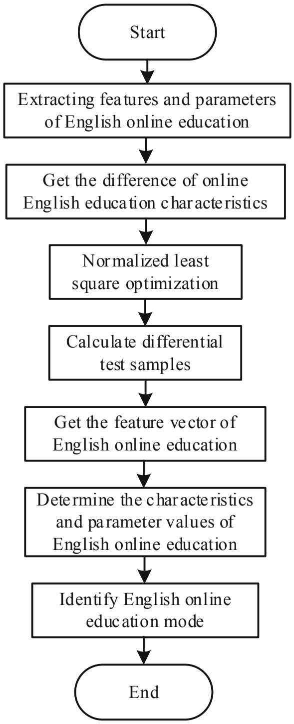

Fig. 3. Flow chart of online English education pattern recognition based on feature self-learning

3 Experiment and Analysis

In order to verify the pattern recognition method of English online education in this study, a comparative experiment is used to verify the pattern recognition method of English online education in the environment of Ubunt14.04, cuda7.5 and tensorflow. The pattern recognition method of English online education in this study is recorded as experimental group A, and the traditional pattern recognition method of English online education is recorded as experimental group B. This paper determines the formula for calculating the correct rate of pattern recognition type in English online education, and compares two groups of pattern recognition methods from subjective and objective perspectives to identify the effect and correct rate of English online education mode.

3.1 Experimental Preparation

Based on this experiment, the online English education courses of a certain platform are divided into four modes: teaching mode based on collective learning, teaching mode based on individual learning, teaching mode based on group learning, and improvement mode based on asynchronous teaching and self-study guidance. Among them, the teaching mode based on collective learning can also be divided into synchronous teaching and asynchronous teaching. The teaching mode based on group learning can also be divided into discussion learning mode and cooperative learning mode. The improved mode combining asynchronous teaching and self-learning guidance can also be divided into courseware learning, online guidance, online question answering, self-learning guidance and learning evaluation Five. Therefore, in this experiment, the above 13 modes were used as the test objects.

According to the above content, the experimental objects, the determined English online education mode, the Ubunt14.04, cuda7.5, tensorflow experimental environment, the Ubuntu 14.04 lts version of Linux operating platform, and the computer software and hardware specifications are set as Table 2.

Table 2. Computer software and hardware configuration table

Experimental environment	Configuration	Configuration instructions
Software environment	Algorithm programming environment	Method framework
	Operating platform	Linux
Hardware environment	GPU model	Nvidia GTX 750
	GPU memory	2 GB
	Processor model	Intel(R)Core(TM)
	CPU model	i3-2130
	Main frequency	3.40 GHz
	RAM	64 GB

Based on this experiment, choose the type of English online education mode, respectively use this experiment, choose three groups of comparison method, verify the research of English online education pattern recognition method.

3.2 Experimental Results of the First Group

Based on this experiment, the first group of experiments are carried out to verify the pattern recognition method of English online education and its effect on the pattern recognition of English online education. In order to record the results of the experiment, 13 groups of modes were selected and expressed in letters as follows: teaching mode a based on collective learning, teaching mode B based on individual learning, teaching mode C based on group learning, improved mode D Based on asynchronous teaching and self-study guidance, synchronous teaching A1, asynchronous teaching A2 and discussion learning C1. The experimental results are as Table 3.

It can be seen from Table 3 that the experiment B group only identified 9 English online education modes, and four English online education modes were not recognized; while the experiment A group recognized all 13 groups of English online modes correctly. It can be seen that the pattern recognition method of English online education in this study can identify all the patterns of English online education.

Table 3. Comparison of pattern recognition effect in English online education

Method	Mode	Whether to recognize	Mode	Whether to recognize
Experiment group A	a	√	a1	√
	b	√	a2	√
	c	√	c1	√
	d	√	c2	√
	–	–	d1	√
	–	–	d2	√
	–	–	d3	√
	–	–	d4	√
	–	–	d5	√
Experiment group B	a	√	a1	√
	b	√	a2	√
	c	√	c1	×
	d	√	c2	×
	–	–	d1	×
	–	–	d2	√
	–	–	d3	√
	–	–	d4	×
	–	–	d5	√

Note: "√" means to identify the correct English online education model; "×" means to identify the wrong English online education model

3.3 Experimental Results of the Second Group

On the basis of the first set of experiments, conduct the second set of experiments. In this group of experiments, the correct rate of pattern recognition in English online education is introduced. Assuming that the number of correct recognition of English online education patterns is R, and the total number of tested English online education is T, the expression of the correct rate of recognition of English online education patterns is P:

$$P = \frac{R}{T} \tag{9}$$

According to the correct rate formula of image recognition shown in formula (9), the results of image recognition and image classification in the first group and the second group of experiments are counted, and three groups of image recognition methods are compared. For the correct rate of image classification recognition results in the test video, the comparison results are as Table 4.

Table 4. Image recognition accuracy

Number of tests	Experimental group A	Experimental group B
186	98.12%	96%
168	98.92%	96.84%
114	99.8%	97.39%
117	99.8%	97.62%
138	99.95%	97.69%
117	99.86%	97.63%
120	99.7%	97.68%
183	98.6%	96.52%
189	98.1%	95.01%

It can be seen from Table 4 that group B is obviously affected by the number of tests. With the increase of the number of tests, the accuracy of pattern recognition in English online education is decreasing. When the number of online English education pattern recognition reaches the maximum, the correct rate of online English education pattern recognition also drops to 95.01%. Experimental group A is less affected by the number of detection, and the accuracy of image classification and recognition is significantly higher than experimental group B. It can be seen that the pattern recognition method of English online education in this study will not be affected by the number of English online education pattern recognition, and the accuracy of English online education pattern recognition is high.

4 Conclusion

To sum up, considering the characteristics of e-learning, the research on the pattern recognition method of English online education. Based on the feature self-learning algorithm and convolution neural network, the feature extraction model of online English education is constructed, and the basic features and depth features are extracted. Z-transform is introduced to standardize the feature data set of English online education, and Laplace gradient function is used to eliminate the feature errors and missing data of English online education. so as to improve the stability, universality and accuracy of pattern recognition. The paper infers the important parameters of online English education by using fuzzy logic theory. According to the characteristics and parameters of online English education, the differences of online English education are calculated, and all the modes of online English education can be effectively identified. However, the pattern recognition method of English online education in this study only starts from the characteristics and parameters of English online education, and does not consider the inference error of fuzzy logic theory to determine the optimal convolution times of convolution neural network. Therefore, in the future research, we need to further study the pattern recognition method of English online education, determine the inference error of fuzzy logic theory, and reduce it to the minimum, find the optimal convolution times of convolution neural network, and further improve the accuracy of pattern recognition of English online education. With the rapid development of network technology, the application of online platform in English education and teaching has become an inevitable trend in the future. Online learning has been widely used. Online teaching platform provides English online learning, online testing and other functions. The intelligent test paper generation technology involved in online testing is becoming more and more mature. The method of pattern recognition of English online education in this study can identify all the patterns of English online education, and the accuracy of pattern recognition of English online education is high.

References

1. Cai, Y., Wang, Z.: The research method of pattern recognition framework based on the distribution. J. Shandong Agric. Univ. (Nat. Sci. Ed.) **51**(2), 266–268 (2020)
2. Chen, L.: Application of artificial intelligence in pattern recognition. Sci. Technol. Innov. Herald **17**(19), 125–127 (2020)
3. Lu, H.: Machine vision pattern recognition method research. Mod. Inf. Technol. **4**(11), 83–85 (2020)
4. Mou, Z., Zhang, Q., Chen, S.: Learning behavior pattern recognition and teaching intervention based on behavior sequence. Mod. Educ. Technol. **30**(5), 74–80 (2020)
5. Guo, H., Zhao, J., Liu, X., et al.: Comprehensive design experiment for pattern recognition based on Python. Exp. Technol. Manage. **36**(8), 178–181 (2019)
6. Xu, X.: Trajectory distribution pattern adaptive recognition simulation based on matching algorithm. Comput. Simul. **36**(11), 410–413,440 (2019)
7. Zhou, B., Xie, G.: Research on key technologies of artificial intelligence in pattern recognition. Mod. Inf. Technol. **3**(22), 110–111 (2019)
8. Liu, H., Hou, S.: Research and practice on bilingual teaching of postgraduate pattern recognition course. J. Anhui Univ. Technol. (Soc. Sci. Ed.) **36**(1), 67–68 (2019)
9. Zhang, Z., Yang, C.: Identifying two refactoring patterns based on hunk and abstract syntax tree. Intell. Comput. Appl. **9**(3), 146–150 (2019)
10. Wang, H., Liu, G., Duan, J., et al.: Transportation mode detection based on self-learning of features. J. Harbin Eng. Univ. **40**(2), 354–358 (2019)
11. Liu, S., Li, Z., Zhang, Y., et al.: Introduction of key problems in long-distance learning and training. Mob. Netw. Appl. **24**(1), 1–4 (2019)
12. Fu, W., Liu, S., Srivastava, G.: Optimization of big data scheduling in social networks. Entropy **21**(9), 902 (2019)
13. Liu, S., Lu, M., Li, H., et al.: Prediction of gene expression patterns with generalized linear regression model. Front. Genet. **10**, 120 (2019)