ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Innovative Design and Practice of Artificial Intelligence-Assisted Comprehensive Literacy Assessment System for English Major Students

Man Guo[1,*]

[1] Foreign Languages School, Wuhan University of Bioengineering, Wuhan, Hubei, 430415, China

SUMMARY: *In this paper, oriented to the needs of comprehensive literacy assessment for English majors, artificial intelligence technology is used as the main means to propose a comprehensive literacy assessment system for students based on neurocognitive diagnosis. The main modules of the system are user information management, import and weighted summation of grades, certificate statistics, certificate management and query, and students' comprehensive literacy assessment and statistics. The literacy assessment function embeds the one-hot coding of students and exercises on the premise of constructing the prerequisite relationship of knowledge points in the modified Q-matrix, and uses the hidden layer of neural network to simulate the interaction between the two. As a result, a neurocognitive diagnostic model based on the modified Q matrix is formed, which realizes the diagnosis of learners' mastery of knowledge points from the neurocognitive level. After processing the students' comprehensive literacy data using principal component analysis, the PLS-SVM algorithm was designed to extract the features of the input data, and the extracted features were used as the SVM to build a prediction model of students' comprehensive literacy. Selecting the experimental samples and setting the comprehensive literacy evaluation dimensions, the comprehensive literacy prediction model for English majors is built within this framework, and the prediction error rate of the model is <0.600.*

KEYWORDS: *neurocognitive diagnosis; principal component analysis; PLS-SVM; students' comprehensive literacy assessment; modified Q matrix*

1 Introduction

Under the background of globalization and cross-cultural communication, English, as a global lingua franca, English majors must have comprehensive literacy such as professional language ability, cross-cultural communication ability, quality of thinking, and humanistic consciousness [1]. English educators urgently need to start from the teaching reality, fully utilize the advantages of digital technology, vigorously promote the modernization of comprehensive literacy assessment of English majors, and comprehensively improve the quality of university English education.

Students' comprehensive literacy is a comprehensive evaluation of students' abilities in all aspects, which can reflect the overall quality and comprehensive ability of students, and it is of great significance to assess it scientifically and objectively. On the one hand, comprehensive literacy assessment can help students understand their own literacy and deficiencies, targeted to improve their own ability, with a wider range of humanities and social sciences knowledge

*18627898522@163.com
https://doi.org/10.65102/is2026101 Published: 30 April, 2026

and practical ability, to promote the overall development of students, so as to improve the competitiveness of employment [2-4]. On the other hand, the results of the assessment can, to a certain extent, reflect the achievements of the double-high construction and the level and quality of talent cultivation, guide the educational practice, promote the better development of professional education, and further promote the educational reform and the improvement of teaching quality [5-7]. The current method of students' comprehensive literacy assessment is relatively single, and students' academic performance is assessed through examination scores, which only singularly reflect students' ability to do problems and language memorization ability, and it is difficult to examine students' comprehensive literacy. In addition, there is a lack of other assessment methods to comprehensively reflect students' comprehensive quality and ability level, such as the evaluation of students' intercultural quality, the evaluation of students' humanistic literacy, and the evaluation of students' critical thinking, etc. [8-10]. In terms of impact, the assessment results rely too much on academic performance and ignore other indicators, which can lead to students focusing only on the development of test-taking ability and neglecting the cultivation and enhancement of other abilities. At the same time, the lack of clarity and scientificity of some assessment standards, the neglect of process evaluation, and the lack of multi-source data lead to assessment results that are not fair and reasonable [11, 12].

In the context of the digital transformation of education, the application of artificial intelligence (AI) technology is becoming more and more widespread. It has powerful capabilities of data analysis, pattern recognition, machine learning, etc., and is able to efficiently process and deeply mine massive educational data to realize accurate, comprehensive, and dynamic evaluation of students' learning process and results [13-15]. Introducing AI technology into the field of students' comprehensive literacy assessment brings a new opportunity to solve the dilemma of traditional assessment.

Currently, AI technology is more involved in English education assessment for language skills such as composition, speaking, listening, and reading comprehension. Literature [16] introduces an AI-based language assessment tool-Speechace, which is able to provide feedback on students' reports of words per minute and bad pauses, as well as color-coded assessment of students' pronunciation accuracy, and assessment of students' overall speaking ability. Literature [17] showed that the ChatGPT tool's checking results are more likely to help students understand and improve their grammatical skills compared to the English grammar checking performed by the Grammarly and ProWritingAid tools. Literature [18] used AI to create a dynamic assessment model and assessment index system for college English writing ability, which realized accurate writing process evaluation, and the evaluation model increased students' interest and enthusiasm in writing. Literature [19] compared the assessment methods of students' English writing ability based on peer assessment and based on AI review in terms of content, sentence expression, vocabulary use, grammatical and spelling correctness, punctuation and other indicators of the article, and the AI technology showed more objective and specific evaluation results. Literature [20] used large-scale language models in generative AI for interactive listening test tasks, which can more realistically assess students' listening ability in complete conversations by focusing on comprehension and interaction behaviors during listening. Literature [21] utilized an AI platform (Atomic Learning) for listening and reading comprehension receptive testing of narrative text, and the assessment results showed that the assessment differences between the two stemmed from the students' need to assess their comprehension of the material. Literature [22] proposed an interactive AI-based reading comprehension assessment system consisting of three parts: speech-to-text and text-to-speech based on natural language processing, reading comprehension assessment based on sentence and word similarity, and graphical presentation of the assessment results based on a visualization module. Literature [23] integrates natural language processing and support vector machine to propose a comprehensive language proficiency assessment method for students, which has an assessment accuracy of up to 98% and realizes automated, objective and scalable assessment results.

In terms of intercultural competence assessment, literature [24] reported formative and summative assessments of intercultural understanding and its communication skills, and the assessment data covered students' written assignments, oral tests, group discussions, cultural projects, self-reflection, peer feedback, and teacher-student exchanges, etc. However, these methods of assessment suffer from instrumental bias, fairness, and an imbalance between linguistic competence and cultural understanding. Literature [25] analyzed intercultural competence dimensions and assessment scales for college students, as well as mathematical model assessment methods, which are simple and easy to operate, and the assessment covers indicators of foreign cultural knowledge, attitudes, and intercultural communication skills. However, this kind of assessment is mostly used in static process and has a long feedback period. Literature [26] integrates Apriori algorithm, SimRank algorithm, and Mann-Kendall method for capturing the relationship between teaching and intercultural competence, evaluating the similarity of students' performance, and segmenting based on the characteristics of intercultural competence, respectively, so as to construct a fuzzy comprehensive assessment model for evaluating students' intercultural competence. Literature [27] constructed a Hidden Markov Model based on Random Forest for extracting linguistic and probabilistic transition patterns of students' language and behaviors in a time series and classifying these patterns, which is able to achieve the classification of intercultural competence.

In terms of thinking skills, literature [28] extracts and identifies student critical thinking indicators from student-created texts with the help of AI and machine learning algorithms in order to assess students' critical thinking skills in analyzing, interpreting, and evaluating information in various contexts. Literature [29] reviewed to three critical thinking and creativity AI assessment tools, i.e., a short-answer scoring tool based on BERT (Bidirectional Encoder Representations from Transformers), a creativity assessment platform based on deep learning, and a scoring rubric for mock interviews based on GPT-3.5. Literature [30] compares student critical thinking skills assessment methods, i.e., ChatGPT-based method and instructor-based method, where the former possesses a wide range of contextual feedback capabilities while the latter possesses a higher degree of accuracy and adherence to scoring criteria. It can be seen that there is room for improvement in the methods of assessing students' critical thinking and creativity skills based on AI technology, and that spoken language is combined with dual assessment by teachers to ensure accuracy and inclusiveness in the assessment results.

AI technology-driven assessment model for students' comprehensive literacy. Literature [31] used data mining technology to establish a comprehensive and objective evaluation system and evaluation model for students' comprehensive literacy, considering multidimensional data such as students' academics, sports, arts, socialization, etc., but the accuracy of the model was stable at 68%-72%. Literature [32] constructed a multi-task neural network model through big data and machine learning algorithms, which was used for students' comprehensive quality evaluation, and also considered students' health status in the evaluation system, and the model also predicted the level of students' comprehensive quality. Literature [33] improves the hierarchical analysis method and introduces the back-propagation neural network to construct an intelligent assessment model for students' comprehensive quality, and the accuracy rate of the assessment is more than 90% by calculating the weights of the students' comprehensive quality indicators and capturing the multidimensional ability and development potential of the students.

In this paper, we first design the functional modules of the comprehensive literacy assessment system for English majors. It proposes a neurocognitive diagnostic model based on modified Q matrix, analyzes the working methods and mathematical expressions of its embedding stage, information aggregation and updating stage, and prediction stage, and sets up a comprehensive literacy assessment system for students based on neurocognitive diagnosis. Secondly, the calculation steps of principal component analysis method are sorted out as the feature extraction method of students' comprehensive literacy data. PLS-SVM method is proposed to plan the construction scheme of comprehensive literacy prediction model for English majors. The experimental samples are selected again, and the distribution of interaction behavior data and comprehensive literacy of the samples are sorted out to demonstrate their feasibility in the research and analysis. Select the assessment dimensions and indicators of comprehensive literacy, carry out principal component analysis, and determine the final structure. Finally, the fitting effect of the diagnostic value and the actual value is compared to verify the reliability of the neurocognitive diagnostic model. Setting the evaluation level of comprehensive literacy indicators, conducting principal component analysis and clustering on the evaluation level, combining with PLS-SVM algorithm to establish a prediction model of comprehensive literacy for English majors.

2 Neurocognitive Diagnostic-Based Measurement of Students' General Literacy

2.1 Basic system architecture

When teachers make assessments, according to the various levels corresponding to different characteristics, it is usually not possible to evaluate quality competence only through theoretical knowledge, and it is also necessary to make judgments based on each teacher's educational practice. Then, the assessment of college students' comprehensive literacy is easily affected by subjective factors, and thus has no objectivity. At this time, it is necessary to make use of computer technology, using fair and objective assessment system of comprehensive literacy of college students to make a more objective and authoritative assessment.

The system framework of the student comprehensive literacy assessment system is shown in Figure 1.

The system is divided into several functional areas, such as user information management, import and weighted sum of grades, certificate statistics, certificate management and query, and student comprehensive literacy assessment and statistics. In order to meet the accuracy and efficiency of comprehensive quality assessment of English majors, the system can import and weighted sum the students' grades. It can also classify and count the types of awards won in various competitions, and meticulously classify and count the types of certificates, such as qualification certificates, competition award certificates and other honorary certificates. It is also possible to assess the “total content” of all the certificates of students, weighted and summed according to certain rules, the results of which can be used as one of the indicators for the evaluation of students' merits and awards, and at the same time, it is also possible to individually count the “content” of a certain type of certificate of students.

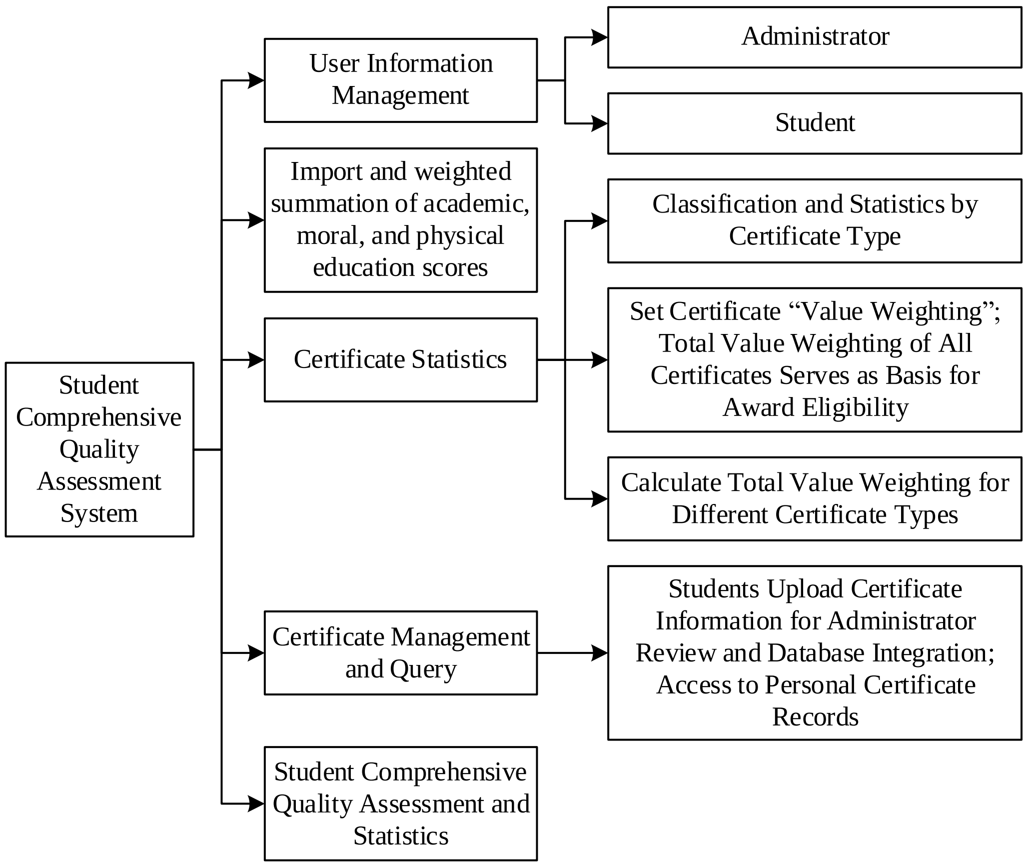


Figure 1: The system framework for the comprehensive quality assessment of students

2.2 Neurocognitive diagnostic modeling framework based on modified Q-matrix

The structure of the neurocognitive diagnostic model based on the modified Q matrix consists of three main parts: the embedding stage, the information aggregation and updating stage, and the cognitive diagnostic result prediction stage. The structure of the model is shown in Figure 2.

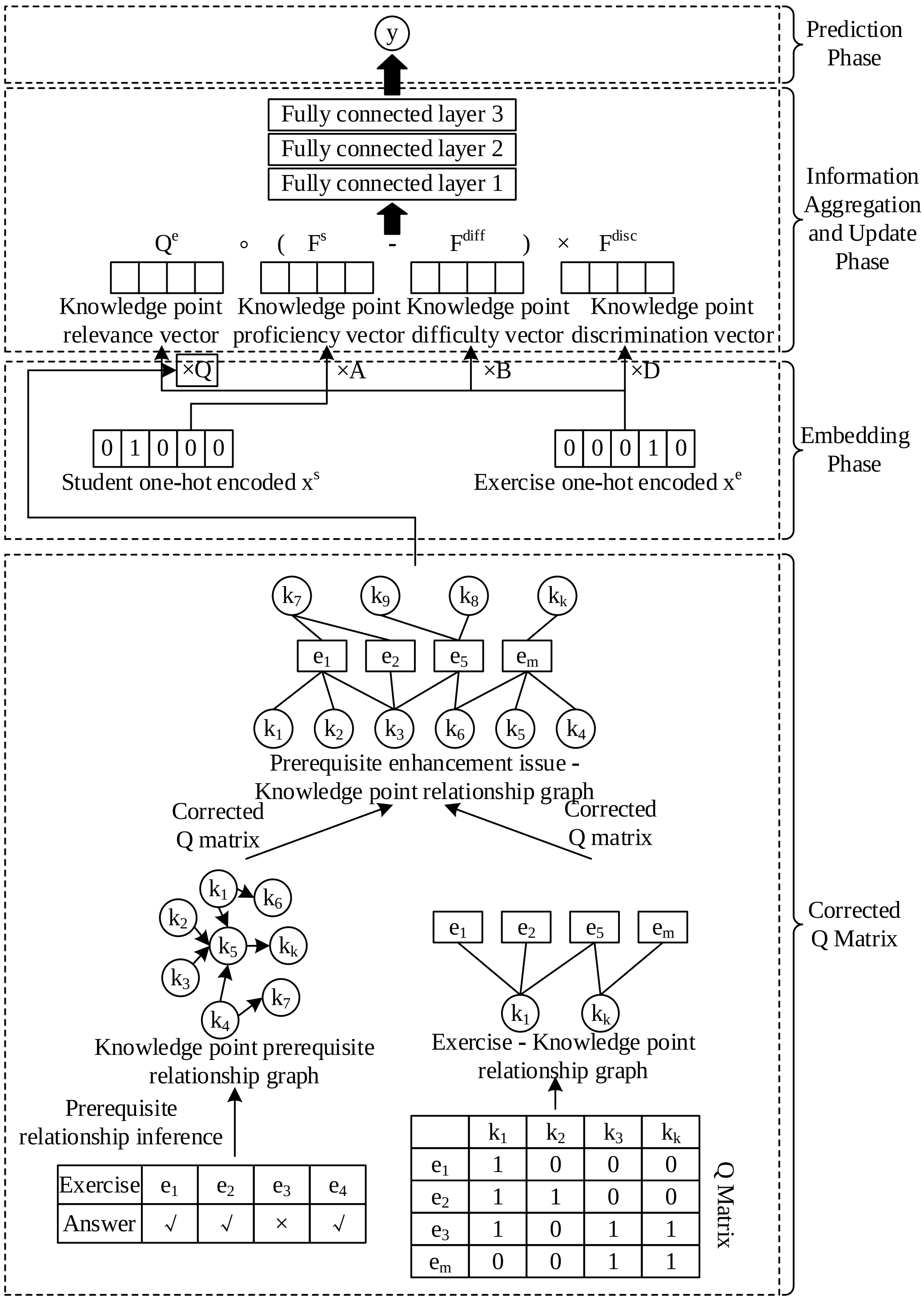


Figure 2: Neural network cognitive diagnosis model based on modified Q matrix

Firstly, the one-hot coding of the student and the one-hot coding of the exercise are embedded together, then each vector is mapped to the formula in combination with IRT for interpretable representation, and the interpretation result is used as the input of the neural network, and finally the student's score is predicted after the training of the three fully-connected layers, which in turn analyzes the state of the student's knowledge.

2.2.1 Embedding phase

In the embedding stage, for the student factor, the student's one-hot encoding x^s is embedded and multiplied with the trained A matrix to obtain the student's proficiency vector F^s, each entry in F^s is a vector of consecutive and taking values in the interval $\left[0,1\right]$ with $F^s = [0.7, 0.1]$ denoting that the student has a higher mastery of the the first knowledge concept to a higher degree and the second knowledge concept to a lower degree[5].

For the practice factor, the one-hot encoding x^e of the practice is embedded and multiplied with the trained B matrix, D matrix and the expert labeled Q matrix to obtain the knowledge point difficulty vector F^{diff}, the differentiation vector of the practice F^{disc} and the knowledge point relevance vector Q^e, respectively.

The mathematical expressions are shown in Eqs. (1)-(4):

$$F^s = sigmoid\left(x^s \times A\right), A \in \mathbb{R}^{N \times K} \tag{1}$$

$$F^{diff} = sigmoid\left(x^e \times B\right), B \in \mathbb{R}^{M \times K} \tag{2}$$

$$F^{disc} = sigmoid\left(x^e \times D\right), D \in \mathbb{R}^{M \times 1} \tag{3}$$

$$Q^e = sigmoid\left(x^e \times \hat{Q}\right) \tag{4}$$

where A, B, and D are trainable matrices and $sigmoid\left(\cdot\right)$ is the activation function.

2.2.2 Information aggregation and updating phase

In the information aggregation and updating phase, based on Item Response Theory (IRT), this model is inspired by Multidimensional Item Response Theory (MIRT) to obtain the input x of the neural network. Its mathematical expression is shown in equation (5):

$$x = Q^e \circ \left(F^s - F^{diff}\right) \circ F^{disc} \tag{5}$$

Three fully connected layers are utilized and the fully connected layer formulas are shown in equations (6)-(8):

$$f_1 = \phi\left(W_1 \times x^T + b_1\right) \tag{6}$$

$$f_2 = \phi\left(W_2 \times f_1 + b_2\right) \tag{7}$$

$$y = \phi\left(W_3 \times f_2 + b_3\right) \tag{8}$$

where $\phi(\cdot)$ denotes the activation function, which is used to compute the output of the hidden layer neurons in the neural network, and the model in this chapter adopts the sigmoid activation function as in Eq. (9).

$$y = \frac{1}{1+e^{-x}} \tag{9}$$

This function is a logistic function with nonlinear characteristics, which is used to map the variables in the model to the $(0,1)$ interval, which is better for the task of binary classification items when the data characteristics are not very different, and the probability model of this function is also known as the "S-curve" because of its graphical characteristics, which is shown in Fig. 3. W_1, W_2, W_3 denote the weights, and b_1, b_2, b_3 denote the biases. In order to make sure that the higher the students' mastery of the knowledge points, the higher the correctness in answering the questions, the model introduces the assumption of monotonicity, and the adopted method is to set W_1, W_2, W_3 each element is restricted to positive values.

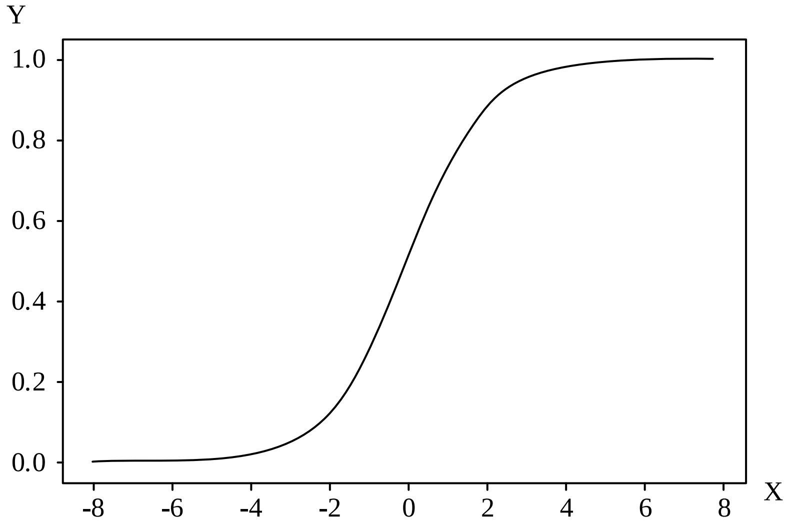


Figure 3: Sigmoid activation function

2.2.3 Forecasting phase

The loss function for the output y and the label r of the real data is shown in equation (10):

$$loss = -\sum_i \left(r_i \log y_i + (1-r_i)\log(1-y_i) \right) \tag{10}$$

After training, the value of F^s is the diagnostic obtained, which indicates the student's knowledge proficiency.

3 Integrated Literacy Prediction Based on Feature Extraction

3.1 Principal Component Analysis Algorithm

In practice, given n samples, the sample data matrix for m indicators is shown in equation (11):

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = \left(X_1, X_2, \ldots, X_p \right) \tag{11}$$

where $X_i = \left(x_{1i}, x_{2i} \ldots x_{ni} \right)'$, $i = 1, 2 \ldots p$.

Because usually the factors are correlated with each other, and increase the intricate interrelationships within the sample, and the principal component analysis can exactly solve this problem, the principal component analysis method can find a few composite indexes $y_1, y_2, \ldots y_n$ to reflect the relationship between the corresponding $x_1, x_2, \ldots x_n$, and $y_1, y_2, \ldots y_n (r \le m)$ are linearly independent, so that the purpose of simplifying the internal structure of matrix X can be achieved.

If you want to calculate the principal components, you need to do the following:

(1) From the known $X_{n \times p}$ sample data matrix, compute $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)' = \frac{1}{n} X' \lambda$;

(2) Calculate the corresponding covariance matrix as in equation (12):

$$V = \frac{1}{n-1} \left(X' - \frac{1}{n} X' \lambda \right) \left(X' - \frac{1}{n} X' \lambda \right)' = \frac{1}{n-1} (S_{ij}) \approx (\sigma_{ij}) \tag{12}$$

where $S_{ij} = \sum_{a=1}^{n} (x_{ai} - \bar{x}_i)(x_{aj} - \bar{x}_j)$.

(3) Normalize the original matrix, i.e., $\tilde{x}_{ij} = \frac{x_{ij} - \tilde{x}_j}{\sqrt{\sigma_{ij}}}$, to form the correlation matrix $R = X'X$ which can also be written as $X = (\tilde{x}_{ij})$;

(4) Using Jacobi's method, find the eigenvalues of R with $\lambda_1 = \lambda_2 = \lambda_3 = \ldots \lambda_p = 0$ and the corresponding eigenvectors $\gamma_1, \gamma_2, \ldots, \gamma_n$, then we can derive the r principal components of X, $Y = \Gamma' X$, where $\Gamma = (\gamma_1, \gamma_2, \ldots \gamma_n)$, and $X = (x_1, x_2, \ldots x_n)$, $Y = (y_1, y_2, \ldots y_r)$;

(5) Calculate the "composite value Z" for each sample and arrange them into single-indicator ordered samples $Z_{(1)} \ge Z_{(2)} \ge \ldots Z_{(n)}$, where $Z = f_1 y_1 + f_2 y_2 + .. f_k y_k (k \le r)$, where each term $f_i y_i$ is given for each sample on the principal component y_i according to the contribution rate see equation (13) to take the value, for the specific analysis of the problem, usually take k so that the cumulative contribution rate of the first k principal components to reach 70% to 80% can be. If the cumulative contribution rate of the current m principal components is greater than 70%, then the number of principal component eigenvalues is determined to be m, of course, when making $\sum_{i=1}^{k} f_i \ge 0.90$ for the best results.

Where the contribution of principal component F is equation (13):

$$F = \frac{\lambda_i}{\sum_{k=1}^{p} \lambda_k} \quad i = 1, 2, \ldots p \tag{13}$$

where λ_i represents the eigenvalues of each principal component.

Thus the variance contribution ratio corresponding to the first k principal components is the weights to calculate the composite statistic as equation (14):

$$F = \sum_{i=1}^{k} \frac{\lambda_i}{\sum_{i=1}^{k} \lambda_i} \times F_i \tag{14}$$

where λ_i denotes the eigenvalues of each principal component, and F_i denotes the eigenvectors of each principal component, i.e., corresponding to the extraction of the factor variables of each principal component, respectively.

In this way, to obtain p principal components of the original index, only the eigenroots of the covariance array S of the original index and the corresponding standardized orthogonal eigenvectors are required. Generally, the eigenvalues and eigenvectors are taken when the cumulative contribution rate is above 85% as the best.

3.2 Implementation of the PLS-SVM algorithm

3.2.1 Modeling Principles

Both Partial Least Squares and Principal Component Analysis methods can be used for feature extraction of data.PCA only performs a bilinear decomposition of the input data X and requires that the variance of the score vector obtained from the decomposition be maximized without considering the effect of the output data Y. In contrast, with PLS method, it is necessary to bilinearly decompose the input X and the output Y at the same time and maximize the correlation between the score vectors obtained from the decomposition of X and the score vectors obtained from the decomposition of Y. Therefore, the features (scores) extracted by PLS method have both input X and output Y information, which is more reflective of the intrinsic connection between inputs and outputs than the PCA decomposition method.

The score matrix T is formed by the score vectors obtained from feature extraction, which is used for SVM modeling instead of the original input X. The dimensionality of the input variables of the SVM model can be greatly reduced, which improves the modeling speed and eliminates the noise introduced by the covariance of high-dimensional variables. The PLS-SVM method is based on such an idea, the principle of which is shown in Fig. 4, where X_{train} and X_{test} represent the input matrices of the training samples and the test samples, respectively, T_{train} and T_{test} represent the score matrices of the training samples and the test samples after feature extraction, respectively, Y_{train} is the known output value of the training samples, and $Y_{predict}$ is the model prediction value of the test samples.

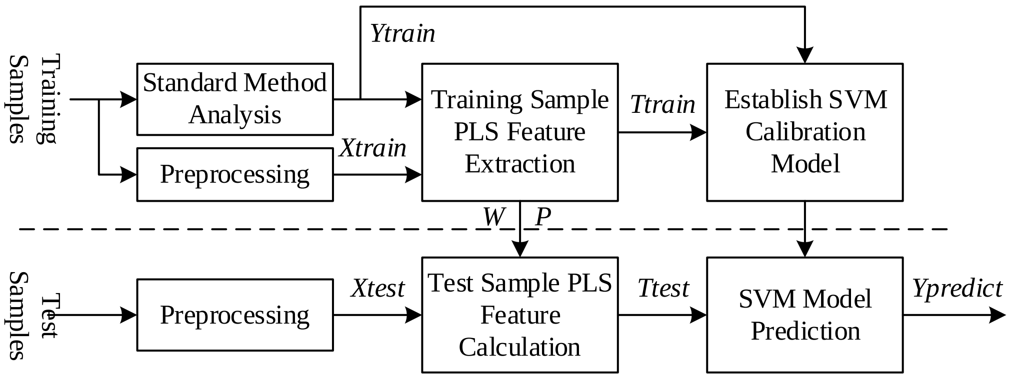


Figure 4: The working principle of PLS-SVM

The processing of PLS-SVM is divided into the following three parts:

(1) Feature extraction: extract the features (score matrix T_{train}) of the training samples using PLS method;

(2) Modeling: train the SVM model with the score matrix T_{train} of the training samples to establish the nonlinear mapping between the score T and the output Y. The SVM model in this paper is built by LS-SVM algorithm, and in fact, it is also possible to build the calibration model by other SVM algorithms;

(3) Prediction: first calculate the score matrix T_{test} of the test sample, and then input T_{test} into the SVM model to get the prediction result $Y_{predict}$.

3.2.2 Algorithmic steps

Taking the single-objective variable problem as an example, the steps of the PLS-SVM algorithm combining PLS1 and LS-SVM are as follows (the input X and output y should be normalized by the standard normalization process before calculation):

(1) PLS feature extraction:

1) Calculate the correlation coefficient vector of the training samples as in equation (15):

$$w_i = X_{train}^T y_{train} / (y_{train}^T y_{train}) \tag{15}$$

2) Normalize the vector of correlation coefficients as in equation (16):

$$w_i = w_i / \|w_i\| \tag{16}$$

3) Calculate the score vector as in equation (17):

$$t_i = X_{train} w_i \tag{17}$$

4) The regression (t on y) is as in equation (18):

$$v_i = (t_i^T y_{train}) / (t_i^T t_i) \tag{18}$$

5) Calculate the X load vector as in equation (19):

$$p_i = X_{train}^T t_i / (t_i^T t_i) \tag{19}$$

6) Calculate the residuals as in equation (20):

$$X_{train} = X_{train} - t_i p_i^T , y_{train} = y_{train} - v_i t_i \tag{20}$$

7) Return to 1) to start the next principal component computation until all the required principal components are obtained, save the vectors t_i, p_i and w_i obtained in each iteration, $i = 1, 2, ..., n$, n is the principal component;

8) The vectors t_i, p_i and w_i obtained from the above computation form the score matrix of the training samples $T_{train} = [t_1, ..., t_n]$, and the loading matrix $P = [p_1, ..., p_n]$ and a matrix of correlation coefficients $W = [w_1, ..., w_n]$;

(2) SVM modeling:

9) Train the SVM model with T_{train}, y_{train}, here the LS-SVM method is used for modeling to solve the system of linear equations in Eq. (21):

$$\begin{bmatrix} 0 & I^T \\ I & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y_{train} \end{bmatrix} \tag{21}$$

where $\Omega = \{\Omega_{kl} \mid k, l = 1, ..., N\}$, $\Omega_{kl} = \varphi(t_k)^T \varphi(t_l) = K(t_k, t_l)$; $k, l = 1, ..., N$. The kernel function $K(t_k, t_l) = \exp\left(-\|t_k - t_l\|^2 / \sigma^2\right)$, t_k and t_l are vectors of training sample scores T_{train}, and the model coefficients α, b are solved.

(3) Model prediction:

10) Calculate the score of the test sample from W, P as in equation (22):

$$T_{test} = X_{test} W (P^T W)^{-1} \tag{22}$$

11) Calculate the predicted value of the test sample as in equation (23):

$$y_{predict}(t) = \sum_{k=1}^{N} \alpha_k K(t_k, t) + b \tag{23}$$

t_k is a vector of scores for the training samples and t is a vector of scores T_{test} for the test samples.

4 Construction of a Predictive Model of Comprehensive Literacy for English Majors

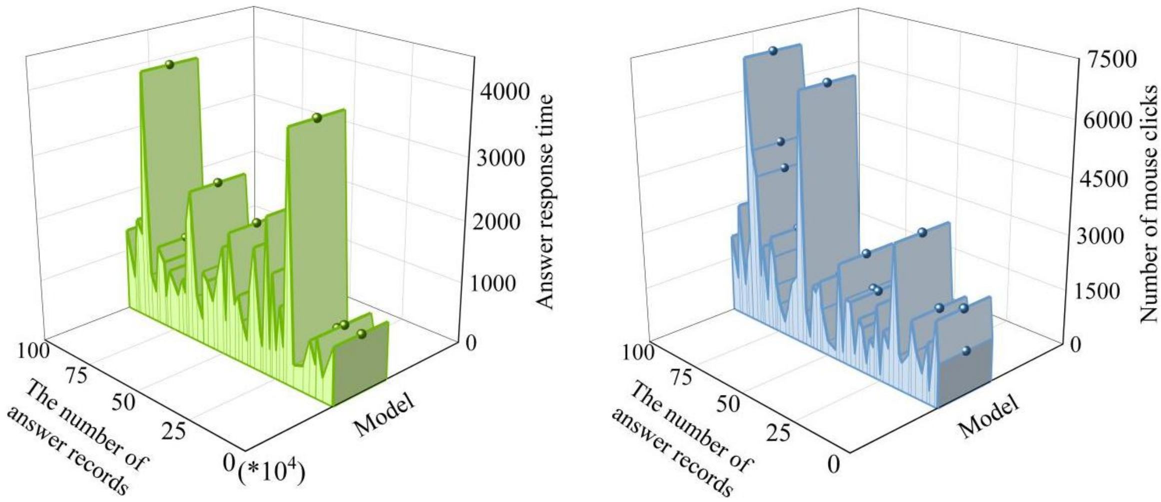
Combined with the existing research, this paper divides the comprehensive literacy of English majors into a total of four dimensions, namely, (H1) language proficiency, (H2) intercultural cognition, (H3) critical thinking ability and (H4) independent learning ability, and makes an attempt to establish a prediction model of comprehensive literacy. A questionnaire was formed based on the four dimensions of (H1) language proficiency, (H2) intercultural cognition, (H3) critical thinking ability and (H4) independent learning ability, in which seven questions (out of 10) were set up under each dimension, and the questions were numbered according to 1-7 according to the dimensions in which they were set up, e.g., the questions under the dimension of (H1) language proficiency were numbered as H11-H17.

A total of 412 second-year English majors in college I were selected as the experimental sample, 412 questionnaires were distributed, 404 valid questionnaires were collected, and the validity rate of the questionnaires was 98.06%. Based on the questionnaire data, the experimental sample data set was formed and the study was analyzed.

4.1 Description of basic information about the sample

4.1.1 Interaction Behavior Data Distribution

Figure 5 shows the distribution of response time and the number of mouse clicks of the experimental samples in the final exam, it can be seen that the distribution of the two types of data are more stable, the response time is mainly distributed within 0~1000s, and the number of mouse clicks is mainly distributed in 0~1000 times. The distribution of the two types of data are more stable, the reaction time is concentrated in the (0,1200) interval, and the number of mouse clicks is distributed in the (0,900) interval, which is in line with the normal situation, and is able to do further research and analysis.



(a) Answer response time (b) Number of mouse clicks

Figure 5: Distribution of interaction behavior data

4.1.2 Integrated literacy performance

Figure 6 visualizes the existing distribution of (H1) language proficiency, (H2) intercultural awareness, (H3) critical thinking ability and (H4) independent learning ability of the experimental sample. It can be seen that the distribution patterns of the four dimensions of the samples are relatively similar, with an overall score range of (3.5,7.5], and most of the distributions are in the range of (5,6), which indicates that the performance of the sample students in the four dimensions has a certain degree of homogeneity and no significant differences.

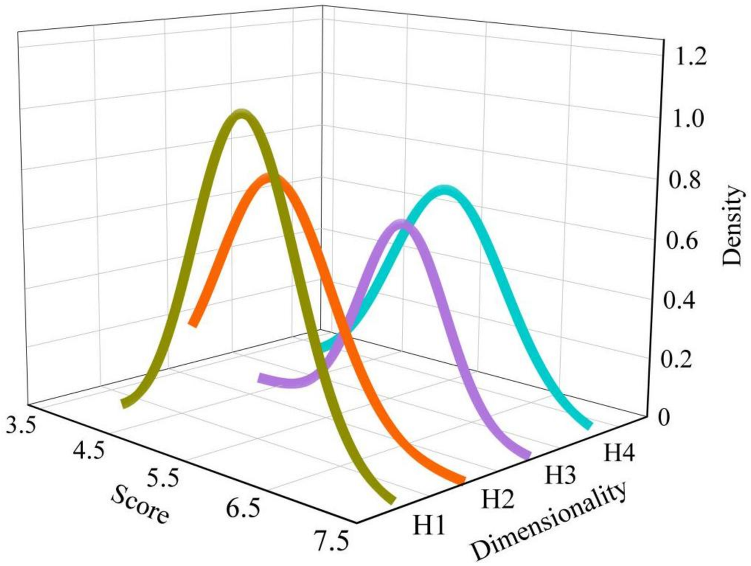


Figure 6: The comprehensive English professional quality performance of the sample

4.2 Indicators for assessing comprehensive student literacy

Each question in the questionnaire was used as an assessment indicator, and the common factors were extracted using principal component analysis with the criterion of eigenroot > 1. The final rotation converged after 13 iterations. The factor loading coefficients and the common values of each assessment indicator after rotation are shown in Table 1.The selection criterion of factor loading is 0.400, among which there are three indicators, H27, H37, and H46, with factor loading <0.4, which indicates that there is no strong correlation between these three indicators and the four factors, and therefore they are deleted. The factor loadings of the other 25 indicators are all above 0.4, indicating that they can cover the required amount of information and express it effectively, so there is no need to correct them. Among them, the loadings of Factor 1 are distributed on indicators H11-H17, which represent the dimension (H1) linguistic competence. The loadings of Factor 2 are distributed on the indicators H21-H26 and represent the dimension (H2) intercultural awareness. Factor 3 loadings were distributed on indicators H31-H36, representing dimension (H3) Critical Thinking Skills. The loadings of Factor 4 are distributed on indicators H41-H45 and H47, representing dimension (H4) independent learning ability.

Table 1: The component matrix after rotation

Serial number	Ingredient			
	Factor 1	Factor 2	Factor 3	Factor 4
H11	0.749			
H12	0.515			
H13	0.731			
H14	0.602			
H15	0.561			
H16	0.526			
H17	0.682			
H21		0.532		
H22		0.538		
H23		0.747		
H24		0.707		
H25		0.751		
H26		0.723		
H27		0.332		
H31			0.579	
H32			0.598	
H33			0.726	
H34			0.678	
H35			0.699	
H36			0.763	
H37			0.304	
H41				0.505
H42				0.774
H43				0.532
H44				0.684
H45				0.519
H46				0.325
H47				0.755

The principal component analysis is still used to extract the common factors with the criterion of eigenroot>1. The final rotation converges after 10 iterations, and the total variance interpretation of the rotated evaluation indexes is shown in Table 2, and the modified rotated component matrices are shown in Table 3.Comprehensively, Table 2 and Table 3 show that the modified index data are still able to extract the 4 common factors, and their total variance interpretation rate reaches 61.711%, and H1, H2, H3, H4 subordinate question items can correspond to the four factors respectively, indicating that the overall extraction effect of the modified factor analysis is better, and can conform to the original evaluation index system.

Table 2: Explanation of the total variance of the rotated evaluation indicators

Ingredient		1	2	3	4	5
Initial eigenvalue	Aggregate	6.575	1.473	1.146	1.053	0.722
	Variance percentage	35.676	10.973	8.521	6.541	1.023
	Cumulative (%)	35.676	46.649	55.17	61.711	62.734
Extract the sum of squares of the loads	Aggregate	6.575	1.473	1.146	1.053	
	Variance percentage	35.676	10.973	8.521	6.541	
	Cumulative (%)	35.676	46.649	55.17	61.711	
The sum of the squares of the rotational load	Aggregate	3.328	2.850	2.697	1.530	
	Variance percentage	18.594	15.203	14.439	13.475	
	Cumulative (%)	18.594	33.797	48.236	61.711	

Table 3: The corrected and rotated component matrix

Serial number	Ingredient			
	Factor 1	Factor 2	Factor 3	Factor 4
H11	0.559			
H12	0.603			
H13	0.522			
H14	0.515			
H15	0.568			
H16	0.531			
H17	0.695			
H21		0.709		
H22		0.625		
H23		0.757		
H24		0.796		
H25		0.698		
H26		0.594		
H31			0.606	
H32			0.677	
H33			0.749	
H34			0.783	
H35			0.59	
H36			0.673	
H41				0.797
H42				0.667
H43				0.607
H44				0.584
H45				0.623
H47				0.769

4.3 Integrated Literacy Prediction Model

4.3.1 Diagnosis of the degree of mastery of knowledge points

The neurocognitive diagnostic model based on modified Q matrix was used for the diagnosis of English knowledge point mastery of the experimental samples, and the true values of the knowledge point mastery of 25 students were randomly intercepted and fitted to the diagnostic values. The fitting effect is shown in Figure 7, and the color-filled area in the figure shows the difference between the real value and the fitted value. It can be seen that the trend of the fitted value and the real value is basically the same, the difference is small, the fitting effect is good, and the neurocognitive diagnostic model is more accurate in the diagnosis of students' knowledge point mastery.

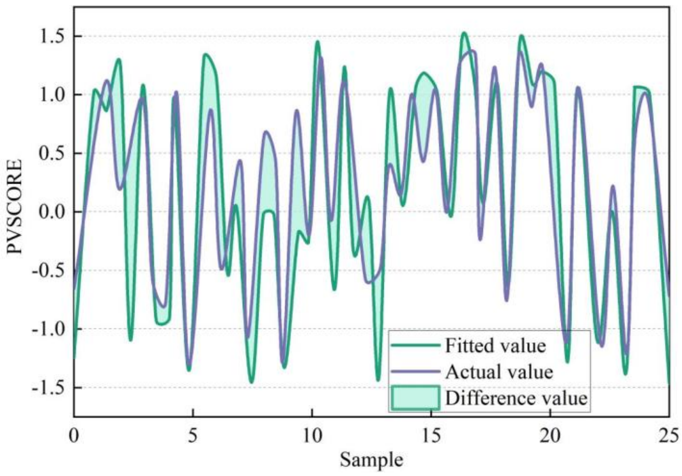


Figure 7: Sample fitting effect based on non-negative least squares regression

4.3.2 Construction of the prediction model

The general literacy of English majors was categorized into the following five grades: (J1) Excellent, (J2) Good, (J3) Moderate, (J4) Pass, and (J5) Fail. Organizing the dataset and performing the principal component analysis on the excellent and good grades produced the same cumulative contribution of 61.711% when four components were obtained. Taking the first principal component as an example, the relationship between (J1) excellent and (J2) good grades in the first principal component is shown in Fig. 8. It can be seen that overall the different grades are differentiated in their performance, but there is still intermingling. That is, in the actual measurement and prediction of comprehensive literacy, students' multidimensional performance should be evaluated comprehensively.

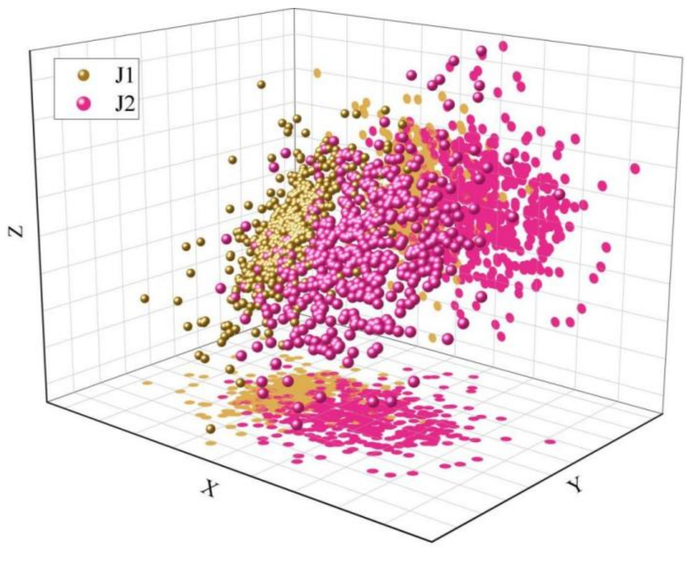


Figure 8: The relationship between J1 and J2

Based on the experimental sample dataset and PLS-SVM algorithm for cross-validation of parameters, after adjusting the range of values of gamma and cost for many times, the range of the final gamma is determined to be [0,0.1], and it grows in steps of 0.01, while the range of values of the penalization parameter cost is determined to be [1,10], to establish the prediction model of the comprehensive literacy for English majors. The cross-validation of the prediction model is plotted in Fig. 9, which can be intuitively obtained that the dark purple region contains the optimal parameters, and the prediction error rate under the optimal combination of parameters in this region is <0.600.

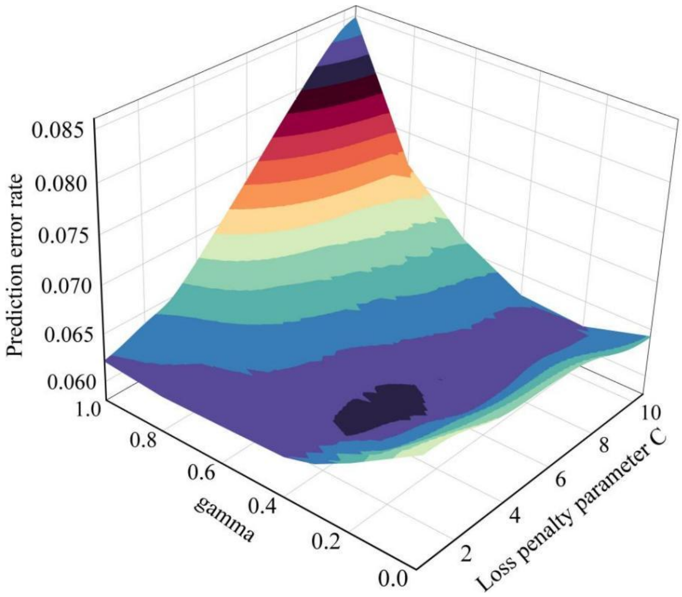


Figure 9: The prediction error rate under different parameter combinations

5 Conclusion

This paper proposes a comprehensive literacy assessment system for students based on neurocognitive diagnosis with the framework of user information management, import and weighted summation of grades, certificate statistics, certificate management and query, and the neurocognitive diagnostic model based on modified Q matrix as the technique for acquiring student performance data. In the diagnosis of the experimental sample knowledge points, the fitted value and the actual value trend are nearly the same, the difference is small, and the fitting effect is good.

Taking (H1) language proficiency, (H2) intercultural cognition, (H3) critical thinking ability and (H4) independent learning ability as the evaluation dimensions of students' comprehensive professional literacy, the prediction model of English majors' comprehensive literacy was established by combining the PLS-SVM algorithm, and the prediction error rate of the model was under 0.600, which had a high prediction accuracy. The comprehensive literacy assessment and prediction method of English majors supported by neural network, PLS-SVM and other artificial intelligence technologies can output assessment and prediction matching with the students' own data, and the validity of assessment and prediction is high, which is expected to be put into the assessment and cultivation of English majors' literacy.

Funding

2025 Industry-University Cooperative Education Project of the Ministry of Education, Sponsored by Shanghai Foreign Language Education Press Information Technology Co., Ltd.: *Digital and Intelligent Teaching Competence Training for ESP Teachers in Application-Oriented Universities Project* （*Project Number:*250502193075022）

References

[1] Phuong, H. Y., & Hua, H. H. (2020). Employers' requirements for English-majored students: A case study from Vietnam. Journal of Contemporary Social Sciences and Humanities, 7(2), 23-32.

[2] Wu, Z., Kakham, P., Mingmit, N., Pholngam, P., & Jujia, C. (2024). Development of Comprehensive Quality Evaluation System Strategies for Higher Vocational College Students in Guangdong. International Journal of Education and Literacy Studies, 12(4), 50-58.

[3] Yiyong, C. H. E. N., SOMJAI, S., & SOMJAI, A. (2024). THE INFLUENCE OF COMPREHENSIVE QUALITY OF STUDENTS AND MANAGEMENT INNOVATION THAT AFFECTS THE PERFORMANCE OF STUDENTS IN FUJIAN VOCATIONAL COLLEGE. Procedia of Multidisciplinary Research, 2(11), 40-40.

[4] Li, Y., & Wang, X. (2021, August). The algorithm and implementation of college English teaching comprehensive ability evaluation system. In EAI International Conference, BigIoT-EDU (pp. 216-224). Cham: Springer International Publishing.

[5] Stepanova, M. (2016). Comprehensive quality assessment in higher education. English Studies at NBU, 2(2), 77-88.

[6] Zhang, X. (2023). Exploration and analysis on the implementation path of high school students' comprehensive quality evaluation under the background of the new college entrance examination. Advances in Educational Technology and Psychology, 7(8).

[7] Yan, L., Sun, H., & Li, N. (2019). Research on Comprehensive Evaluation of College Students Based on Improved Analytic Hierarchy Process. Innovative Education Research, 7(3), 275-281.

[8] Su, C., & Zhang, J. (2024). The Advice on Assessment System for College Student Comprehensive Quality. Advances in Social Development and Education Research, 1(3), 166-169.

[9] Zhang, L. (2020). Comprehensive analysis of students' performance based on student comprehensive analysis model. Academic Journal of Computing & Information Science, 3(4), 66-72.

[10] Lu, W., Jia, C., & Zuo, J. (2021). Application of Fuzzy Comprehensive Evaluation in Comprehensive Quality Evaluation of Higher Education Students. International Journal of Emerging Technologies in Learning, 16(12).

[11] Xu, Z. (2022). Research on the comprehensive Literacy test and evaluation of English course students based on cloud education big data. Advances in Economics and Management Research, 3(1), 441-441.

[12] Wang, Z. (2024). Evaluation method for the comprehensive quality of students based on multi-source data fusion. Asia Pacific Education Review, 1-8.

[13] Divasón, J., Martínez-de-Pisón, F. J., Romero, A., & Sáenz-de-Cabezón, E. (2023). Artificial intelligence models for assessing the evaluation process of complex student projects. IEEE Transactions on Learning Technologies, 16(5), 694-707.

[14] Rybinski, K., & Kopciuszewska, E. (2021). Will artificial intelligence revolutionise the student evaluation of teaching? A big data study of 1.6 million student reviews. Assessment & Evaluation in Higher Education, 46(7), 1127-1139.

[15] Ucheagwu-Okoye, O. M. (2025). The Role of Artificial Intelligence (AI) in Student Assessment and Evaluation in Educational Leadership in Tertiary Institutions in Anambra State. Unizik Journal of Educational Laws and Leadership Studies, 1(1).

[16] Ningsih, F. (2024). ANALYZING STUDENTS'ENGLISH-SPEAKING SKILLS USING SPEECHACE: INSIGHTS FROM AN AI-POWERED ASSESSMENT TOOL. Pedagogic Research-Applied Literacy Journal, 1(3), 111-124.

[17] Schmidt-Fajlik, R. (2023). ChatGPT as a grammar checker for Japanese English language learners: A comparison with Grammarly and ProWritingAid. AsiaCALL Online Journal, 14(1), 105-119.

[18] Wu, X. (2022). Dynamic evaluation of college English writing ability based on AI technology. Journal of Intelligent Systems, 31(1), 298-309.

[19] Lu, H. (2025). Effect of peer assessment and artificial intelligence review on English writing and evaluation ability. International Journal of Instruction, 18(4), 517-530.

[20] Runge, A., Attali, Y., LaFlair, G. T., Park, Y., & Church, J. (2024). A generative AI-driven interactive listening assessment task. Frontiers in Artificial Intelligence, 7, 1474019.

[21] Fakhrurriana, R., & Herdina, G. F. (2024, March). Assessment Test of Receptive Skills: Listening and Reading Comprehension of Recount Text Using AI Platfrom (Atomic Learning). In Proceeding International Conference on Religion, Science and Education (Vol. 3, pp. 187-192).

[22] Yu, H., Jeon, E., Oh, S., Son, D., Lee, S., & Lim, K. S. (2022, October). Design of Interactive Reading Comprehension Competence Assessment System using AI. In 2022 13th International Conference on Information and Communication Technology Convergence (ICTC) (pp. 2062-2064). IEEE.

[23] Pathak, P., Pearlin, E., Punithaasree, K. S., Rajeswari, A., & Alphonse, F. R. (2024, December). Automating Language Proficiency Evaluation in English Language Teaching: An NLP Framework. In 2024 International Conference on Communication, Control, and Intelligent Systems (CCIS) (pp. 1-5). IEEE.

[24] Rahmawati, Y. (2023). Assessing cross-cultural understanding and intercultural communication skills in EFL classrooms: Challenges, best practices, and perceptions. NextGen Education Review Journal, 1(2), 22-32.

[25] Long, J., & Lin, J. (2022). An empirical study on cultivating college students' cross-cultural communicative competence based on the artificial-intelligence English-teaching mode. Frontiers in Psychology, 13, 976310.

[26] Zhang, Q., Mohammad Ismail, M. I. R., & Zakaria, A. R. B. (2025). Enhancing intercultural competence in technical higher education through AI-driven frameworks. Scientific Reports, 15(1), 22019.

[27] Niu, T., & Zhu, Y. (2024, September). English Cross-Cultural Competence of College Students Analysis using Hidden Markov Model with Random Forest. In 2024 International Conference on Distributed Systems, Computer Networks and Cybersecurity (ICDSCNC) (pp. 1-5). IEEE.

[28] Poce, A. (2024, June). Using AI for critical thinking assessment: A digital humanities education experience. In The Learning Ideas Conference (pp. 94-102). Cham: Springer Nature Switzerland.

[29] Suazo-Galdames, I. C., & Chaple-Gil, A. M. (2025). AI-Driven Assessment Systems in Higher Education: Effectiveness for Enhancing Critical Thinking and Creativity. Ingénierie des Systèmes d'Information, 30(6).

[30] Trikoili, A., Georgiou, D., Pappa, C. I., & Pittich, D. (2025). Critical Thinking Assessment in Higher Education: A Mixed-Methods Comparative Analysis of AI and Human Evaluator. International Journal of Human–Computer Interaction, 1-14.

[31] Luo, Y. (2024, September). Student Comprehensive Quality Evaluation System Based on Data Mining and Analysis. In Proceedings of the 2024 International Symposium on Artificial Intelligence for Education (pp. 369-374).

[32] Han, Y., Wang, S., & Zhou, L. (2024, May). Evaluation and Prediction Algorithm of Students' Comprehensive Quality Based on Big Data and Machine Learning. In 2024 International Conference on Telecommunications and Power Electronics (TELEPE) (pp. 136-141). IEEE.

[33] Zhang, L. (2025). A student comprehensive quality assessment system based on improved analytic hierarchy process and back propagation neural network. Discover Artificial Intelligence, 5(1), 1-20.