Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 8351806, 10 pages
https://doi.org/10.1155/2022/8351806

*Research Article*

# Application of Unbalanced Data Classification Based on CSD-ELM in English Network Teaching Mode

**Yingmei Xie**

*School of Humanities and Arts, Jangsu Shipping College, Nantong, Jiangsu 226001, China*

Correspondence should be addressed to Yingmei Xie; xie_yingmei@outlook.com

Received 31 December 2021; Revised 13 January 2022; Accepted 18 January 2022; Published 17 February 2022

Academic Editor: Kalidoss Rajakani



In order to accurately predict the frame loss phenomenon, this study proposes a continuous frame loss test for CSD-ELM short-range wireless communication. First, the noise and trend direction of CSD-ELM are removed, the preprocessing of the data is completed, then, these data are injected into the BP neural network as a training sample, continuous frame loss tests are performed on the communication data, the results are output, and the calculation formula is obtained. The test results show that the method has higher accuracy, lower false alarm rate, and shorter test time and is suitable for the detection of continuous data frame loss in short-distance wireless communication. Then, this study proposes an integrated WELM algorithm, which combines the features of AdaBoost, which can generate strong classifiers through the integration of weak classifiers, so as to achieve the best performance for imbalanced data classification. The algorithm assigns different methods to update the weights of different types of samples, so as not to violate the skewness of the weight distribution. The sigmoid function uses error calculation technology, which can separately reflect the classification characteristics of the two samples, and improves the antinoise function of the algorithm. Finally, this study introduces how to analyze and design a college network English learning management system that can meet the requirements of "network system and classroom English teaching mode" in the internet environment. It solves the time limitation of existing English classrooms, expands college English courses to be student centered, creates interactive online classroom spaces, realizes English classrooms regardless of time and region, avoids teachers' duplication of work, improves teachers' teaching efficiency, and ultimately promotes the improvement of students' learning efficiency.

## 1. Introduction

Aiming at the problems that the traditional data frame loss detection method is susceptible to external interference, the detection result is inaccurate, the frame loss phenomenon is untimely, and a CSD-ELM continuous frame loss test method for short-distance wireless communication is proposed. We analyze the characteristics of continuous frame loss of short-range wireless communication CSD-ELM and analyze the reasons [1]. Based on this, we created a noise model and used least squares and average slope methods to remove noise and trend elements from the data. It is used to complete the processing of the CSD-ELM process and completely inject the preprocessed data as training samples into the BP neural network classifier to determine whether frame loss occurs and to perform a frame loss test [2]. The simulation results show that the above method has high detection accuracy, short detection time, effectively improves detection efficiency and can be widely used in the test of data frame loss during transmission. Extreme learning machines based on this are widely used to study classification problems because they are fast, simple, easy to implement, and highly versatile, but the use of extreme learning machines to learn imbalanced data classification started relatively late [3]. Compared with traditional neural networks, the more nodes in the hidden layer are, the greater the influence of random parameters on the output stability [4]. The weight of training samples depends on the input data of the algorithm, which affects the generalization performance of the algorithm, and the antinoise ability of the algorithm needs to be improved. Then, on how to improve the ability of weighted extreme learning machine to classify unbalanced data, three types of weighted extreme learning machine classification algorithms for unbalanced data are proposed.

In the latest research, scholars have analyzed the factors affecting the effectiveness of online English classroom teaching and suggestions for improvement and found that unbalanced data classification can help improve the efficiency of online teaching [5]. The model-building method and the key elements of the online teaching reform confirm the usability of the CSD-ELM method [6].

Finally, this study aims to explore the application and effectiveness of the English teaching network system in college English teaching. The author compared English teaching between the two juniors [7]. Class A uses the English teaching network system, and students complete their homework online; class B is the control group, which uses traditional teaching methods, and the homework format is the same as usual. Except for the different teaching methods, the teaching content and homework of the two classes are the same. After completing the module training, the test will be conducted in the classroom. According to the data, we analyzed the excellent rate and pass rate of the two classes. In this study, students are impressed by what they have learned through the English teaching network system, and their memory speed is also very high. Even students who lack sufficient basic knowledge express that they are willing to accept the online learning method of the English online teaching system. In the same test comparison, class A scored higher and performed better than class B. After the first semester, the excellence rate and passing rate of class A have increased, but the growth of class B is relatively small, which is significantly lower than that of class A.

## 2. Related Work

The literature [8] introduces the research background, methods, and importance of online college English reading education, combined with the reform of Chinese college English education and the demand for training applied professionals, and analyzes the current situation and challenges of Chinese college English reading education. The literature introduced the enhanced ELM for the diagnosis of pleural effusion [9]. This model uses the chaotic initialization mechanism and the differential evolution mechanism to improve the global and local optimization capabilities of the high salvage algorithm. The literature introduces an extreme learning machine (ECSMA-KELM) [10] based on the elite chaos slime mold algorithm. The proposed model uses the elite mechanism and the chaos mechanism to expand the global and local search capabilities of the slime mold algorithm while converting these two capabilities into a relatively stable state. The literature introduces the research background and importance of multilabel learning and category-imbalanced learning [11]. Next, we will explain related methods to deal with multilabel learning and category-imbalanced learning and then combine the conventional treatment of category imbalances in two or more categories. This is another important issue to be solved in this study, involving technology and multiple trademark areas [12]. The literature introduces the PSO particle swarm optimization algorithm to find the best combination of thresholds, and the integrated AdaBoost learning algorithm can find more reasonable label weights to improve the classification performance across multiple labels.

## 3. CSD-ELM and Imbalanced Data Classification Technology

*3.1. ELM Algorithm.* The extreme learning machine (ELM) is a single-layer hidden layer feedback network (SLFN) training method proposed by Professor Huang Guangbin from the Nanyang Technological University in Singapore. The problem is solved by finding the generalized Moore-Penrose inverse matrix instead of directly randomly assigning hidden layer weights and biases, which is contrary to the traditional neural network training method used to continuously correct iterative errors [13]. There is no need to repeat the entire training process, and the solution is unique, so it can greatly reduce the training time while still providing network generalization functions.

Suppose that the training set contains $N$ samples, these samples can be classified into $m$ classes. Here, the $i$-th training set can be expressed as $(x_{\mathrm{i}}, t_i)$, where $x_i$ is an $n$-dimensional input vector and $t_i$ matches. The mathematical model of ELM can be expressed as follows:

$$H\beta = T. \tag{1}$$

ELM is looked at from an optimization perspective. In order to minimize learning errors and improve the versatility of the model, it is necessary to minimize the sum of $\|H\beta - T\|^2$ and $\|\beta\|^2$ at the same time, so the following formula can be used to explain the problem:

$$\begin{cases} \text{Minimize:} & \mathrm{Lp}_{\mathrm{ILM}} = \dfrac{1}{2}\|\beta\|^2 + C\dfrac{1}{2}\sum_{i=1}^{N}\|\xi_i\|^2 \\ \text{Subject to:} & h(x_i)\beta = t_i^T - \xi_i^T,\ i = 1, 2, \dots, N. \end{cases} \tag{2}$$

Regarding the cost-sensitive learning concept, Zong et al. focused on the impact of category imbalance on extreme learning machines. A weighted machine (WELM) with an extreme learning algorithm is proposed. The WELM algorithm optimizes formula (3) into the following form:

$$\begin{cases} \text{Minimize:} \quad \mathrm{Lp}_{\mathrm{ELM}} = \frac{1}{2}\|\beta\|^2 + C\frac{1}{2}W\sum_{i=1}^{N}\|\xi_i\|^2 \\ \text{Subject to:} \quad h(x_i)\beta = t_i^T - \xi_i^T,\ i = 1,\ldots,N. \end{cases} \tag{3}$$

Among them, $W$ is the $N$th order diagonal matrix, and the value of each element on the diagonal corresponds to the penalty factor control parameter of the corresponding sample. The higher the corresponding value, the more important is the sample. Zong et al. provide two weight distribution methods:

$$\begin{aligned} &\text{WELM1:}\ W_{ii} = 1/\#(t_i), \\ &\text{WELM2:}\ W_{ii} = \begin{cases} \frac{0.618}{\#(t_i)}, & \text{if } \#(t_i) > \text{AVG}, \\ \frac{1}{\#(t_i)}, & \text{if } \#(t_i) \le \text{AVG}. \end{cases} \end{aligned} \tag{4}$$

Among them, $\#(t_i)$ represents the number of samples belonging to the category $t_i$ in the training set, and AVG represents the average number of training samples in all categories.

*3.2. Unbalanced Data Classification.* In the sampling method, the imbalance ratio is a very important concept, which determines the number of samples. Therefore, the author proposes a multivariate IRL imbalance metric, which is represented by a set of labeled samples and the current maximum value. It is assumed that the dataset $D$ is labeled as size $q$, as shown in equation (5):

$$\begin{aligned} \mathrm{IRL}(l) &= \frac{\operatorname{argmax}_{l=l_1}^{l_q}\left(\sum_{i=1}^{|\mathrm{DI}} h(l', y_i)\right)}{\sum_{i=1}^{|\mathrm{DI}} h(l', y_i)}, \\ h(l, y_i) &= \begin{cases} 1, & l \in y_i, \\ 0, & l \notin y_i. \end{cases} \end{aligned} \tag{5}$$

At the same time, the author recommends using meanIR to represent the average imbalance of multiple label datasets:

$$\mathrm{meanIR} = \frac{\sum_{l=l_1}^{l_q} \mathrm{IRL}(l)}{q}. \tag{6}$$

It is assumed that data $D$ with multiple labels have a $q$-dimensional label space. $D_j$ represents the dataset of the second label category $y_j$:

$$D_j = \left\{\left(x_i, \phi\left(Y_i, y_j\right)\right) | 1 \le i \le N\right\}, \quad \text{where } \phi\left(Y_i, y_j\right) = \begin{cases} +1, & \text{if } y_j \in Y, \\ -1, & \text{if } y_j \notin Y. \end{cases} \tag{7}$$

$D_{jk}$ is a multitype dataset that combines $(y_j, y_k)$ coupling. $D_{jk}$ is converted into three types of datasets $D_{jk}^{\mathrm{tri}}$ to reduce the increase in imbalance level caused by the binding of markers:

$$\begin{aligned} D_\beta^{ti} &= \left\{\left(x_i, \varphi\left(Y_i, y_j, y_k\right)\right) | 1 \le i \le N\right\}, \\ \varphi\left(Y_i, y_j, y_k\right) &= \begin{cases} 0, & \text{if } y_j \notin Y_i \text{ and } y_k \notin Y_i, \\ +1, & \text{if } y_j \notin Y_i \text{ and } y_k \in Y_i, \\ +2, & \text{if } y_j \in Y_i. \end{cases} \end{aligned} \tag{8}$$

The predicted actual value of $f_j(x)$ in the category label is the sum of the results of the above classifiers:

$$f_j(x) = g_j(+1|x) + \sum_{y_k \in I_k} g_{jk}(+2|x) \tag{9}$$

In order to measure the quality of models with multiple tags, scientists have proposed a set of performance evaluation tables. It can be roughly divided into two types: sample-based grade scale and label-based grade scale.

The sample-based grade scale first grades a test sample and then averages the entire test set as the final result. Such evaluation measures include noise loss, coverage, grade loss, and average accuracy [14]. However, on the issue of category imbalance, this type of rating scale is obviously biased toward most categories, so only a brief introduction is given here.

Based on the mark-based scoring scale, the effect of the classifier on a mark is first measured, then, aggregated, and then marked. Among multiple labels, the minority class is usually written as a positive class, and the majority class is written as a negative class. The effect statistics of this general classifier can be expressed as a confusion matrix.

In Table 1, the diagonal lines TP and TN represent the number of true cases and the number of true negative cases, respectively, indicating that the predicted category matches the true category.

The precision rate is the correct probability of predicting the positive category, and the recall rate is the predicted probability of the true-positive category. Two formulas are as follows:

Table 1: Confusion matrix of classification results.

| | Predicted positive class | Predicted negative class |
|---|---|---|
| True positive | TP (real positive) | FN (false negative) |
| True negative | FP (false positive) | TN (true negative) |

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \qquad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

The precision rate and the recall rate are a pair of contradictory indicators. Generally, a higher precision rate usually results in a lower recall rate, and a higher recall rate usually results in a lower precision rate. It can be combined into a main F-measure indicator to balance the two. It is calculated as follows:

$$F - \text{measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{11}$$

The final result of multiple markers can be obtained by summing the average value in two ways using the performance score of each marker. There are two estimation methods: macroaverage and microaverage [15]. The macroaverage first obtains the classification efficiency of each tag, then sums the classification efficiency of all tags, and the averages it, while microaverage directly combines all tags, the statistical information will be added, and then, the resulting classification result is used as the final result. The formulas for these two estimation methods are as follows:

$$\text{Macro} - \text{average} = \frac{1}{q}\sum_{i=1}^{q} \text{eval}M\left(\text{TP}_i, \text{FP}_i, \text{TN}_i, \text{FN}_i\right),$$
$$\text{Micro} - \text{average}\,\text{eval}M\left(\sum_{i=1}^{q}\text{TP}_i, \sum_{i=1}^{q}\text{FP}_i, \sum_{i=1}^{q}\text{TN}_i, \sum_{i=1}^{q}\text{FN}_i\right). \tag{12}$$

This topic explores the imbalance in multiple category labeling. The combination of F-measure, macroaverage, and microaverage was selected as the entire multimarker scoring scale. For convenience, two evaluation methods are selected. It is called the macroaverage F index (Macro-F) and the microaverage F index (Micro-F).

*3.3. Unbalanced Data Classification Model Based on CSD-ELM.* ELM network structure is not only suitable for single-component learning but also suitable for multicomponent learning. When training with multiple labels, equation (13) is still valid, but the number of output nodes does not represent the number of categories, but the number of labels for data categories represents with multiple labels, that is, $q$-associated output nodes.

$$\begin{cases} \text{Minimize:} \quad \text{Lp}_{\text{ELM}} = \dfrac{1}{2}\|\beta\|^2 + C\dfrac{1}{2}\sum_{i=1}^{N}\|\xi_i\|^2 \\ \text{Subject to:} \quad h(x_i)\beta = t_i^T - \xi_i^T,\ i = 1, 2, \ldots, N. \end{cases} \tag{13}$$

For problems with multiple labels, one sample can be associated with multiple samples at the same time. At this time, the threshold function $th(x)$ must be set, and the category standard is predicted by the following formula:

$$\text{label}(i) = \begin{cases} +1, & \text{when } f_i(x) \geq th(x), \\ -1, & \text{when } f_i(x) < th(x). \end{cases} \tag{14}$$

We need to measure the dimensions to determine whether they are optimal. In this chapter, we will choose two optimization indicators: macro-F or micro-F.

In the process of PSO optimization, each particle will dynamically change its position and speed and learn from its own experience and the experience of other members of the population. The updated method is as follows:

$$\begin{cases} v_{\text{id}}^{k+1} = v_{\text{id}}^{k} + c_1 \times r_1 \times \left(p\text{best} - x_{\text{id}}^{k}\right) + c_2 \times r_2 \times \left(g\text{best} - x_{\text{id}}^{k}\right), \\ x_{\text{id}}^{k+1} = x_{\text{id}}^{k} + v_{\text{id}}^{k+1}. \end{cases} \tag{15}$$

Weighted extreme value learning is an effective algorithm for dealing with category imbalance, but it has two disadvantages. (1) The calculation time increases with the increase in the training sample size; (2) there is no flexibility in error correction. The label-weighted extreme learning machine improves the tolerance to errors of minority instances by increasing the expected output of smaller labels. In addition, since the weight matrix is not optimized, the time complexity is basically the same as that of the standard ELM.

$$\begin{cases} \text{Minimize:} \quad \mathrm{Lp}_{\mathrm{BLM}} = \frac{1}{2}\|\beta\|^2 + C\frac{1}{2}W\sum_{i=1}^{N}\|\xi_i\|^2 \\ \text{Subject to:} \quad h(x_i)\beta = t_i^T - \xi_i^T,\ i = 1, \ldots, N. \end{cases} \tag{16}$$

It is assumed that the expected value matrix $T$ has $m$ rows and $N$ columns. Among them, $m$ represents the number of categories, and $N$ represents the number of samples in the training set. $T$ is expressed as follows:

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1N} \\ t_{21} & t_{22} & \cdots & t_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mN} \end{bmatrix}. \tag{17}$$

To set tag weights, Yu et al. proposed two weight distribution methods for the classification of two and several categories:

$$\mathrm{LW^1ELM}: t_{ij} = \begin{cases} \frac{\#\max(C)}{\#(C_i)}, & \text{if } x_j \in C_i, \\ -1, & \text{if } x_j \notin C_i. \end{cases} \tag{18}$$

As well as:

$$\mathrm{LW^2ELM}: t_{ij} = \begin{cases} \sqrt[2]{\frac{\#\max(C)}{\#(C_i)}}, & \text{if } x_j \in C_i, \\ -1, & \text{if } x_j \notin C_i. \end{cases} \tag{19}$$

At the same time, two weight distribution methods are given as follows:

$$\begin{aligned} \mathrm{LW^{\,1}\,ELM}: t_{ij} &= \begin{cases} \frac{\sim(C)}{(C_i)}, & \text{if } x_j \in C_i, \\ -1, & \text{if } x_j \notin C_i, \end{cases} \\ \mathrm{LW^2ELM}: t_{ij} &= \begin{cases} \sqrt[2]{\frac{\sim\#(C)}{\#(C_i)}}, & \text{if } x_j \in C_i, \\ -1, & \text{if } x_j \notin C_i. \end{cases} \end{aligned} \tag{20}$$

*3.4. Simulation Experiment.* The experiment is basically completed on 12 multilabel test datasets, which cover various scenarios, such as text, speech, and biology. Each dataset has its own sample number, category label number, label density, and imbalance factor. Table 2 shows specific information about these datasets.

The LCard marking power is an indicator to measure the average number of marks in the sample. The ratio of the corresponding marking power to the number of marks is called the LDen marking density, and its formula is as follows:

$$\begin{aligned} \mathrm{LCard}(D) &= \frac{1}{N}\sum_{i=1}^{N}|Y_i|, \\ \mathrm{LDen}(D) &= \frac{1}{q}\cdot \mathrm{LCard}(D). \end{aligned} \tag{21}$$

The hardware environment of the experiment is as follows: Intel Core i7-555U processor, 3.1 GHz main frequency, 8 GB memory, 1 TB hard drive, Windows 8.1 operating system, and Matlab 2015b programming environment.

As mentioned above, when selecting the threshold, there are two targets for selecting the appropriate PSO particle swarm algorithm, macro-F, and micro-F.

Finally, the analysis of experimental parameters is also very important for this model. In this chapter, we will select the dataset scene with less than 10 labels and the cal500 dataset with more than 100 labels and use them as examples to test the corresponding models in macro-F dimensions with different $L$ and $C$ parameters, tested during the entire experiment.

These two datasets have minimum and maximum values within the selected parameter range, and the maximum value is not in the boundary state. In other words, you can see that the parameter range contains the maximum category. This also proves that the parameter range is effective. Therefore, it is recommended to set the parameter range as $L \in \{50, 100, \ldots, 1000\}$ and $C \in \{2^1, 2^2, \ldots, 2^{20}\}$.

We also analyzed the relationship between the number of iterations of the particle swarm algorithm and the number of markers through experiments. Theoretically, the number of markers indicates the size of the marker space, which will increase the search range in the high-dimensional space, and it requires more iterations. The number of convergence iterations in scene and cal500 are 20 and 60 generations, respectively. It can be concluded that the more the markers, the more the iterations.

## 4. Design and Application of English Network Teaching System

### *4.1. Demand Analysis of English Network Teaching System*

#### *4.1.1. Performance Requirements*

*(1) Maintainability.* Maintenance method: The purpose of designing this method is that even if the backup data are damaged or lost, the backup data can be backed up and restored. As a management information system application, the network English reading teaching system of a university gives priority to the maintainability of the design, and the system has a complete and simple automatic data backup and recovery function [16].

Table 2: Dataset description in this chapter.

| Dataset | Number of samples | Feature number | Number of marks | LCard | LDen |
|---|---|---|---|---|---|
| Image | 600 | 294 | 5 | 1.235 | 0.247 |
| Emotions | 593 | 72 | 6 | 1.869 | 0.311 |
| Scene | 2407 | 294 | 6 | 1.074 | 0.179 |
| Flags | 194 | 19 | 7 | 3.392 | 0.485 |
| Yeast | 2417 | 103 | 14 | 4.237 | 0.303 |
| Birds | 645 | 260 | 19 | 1.014 | 0.053 |
| Tinc2007 | 28596 | 500 | 22 | 2.220 | 0.101 |
| Miiflicki’ | 25000 | 150 | 24 | 3.716 | 0.155 |
| Genbase | 662 | 1186 | 27 | 1.252 | 0.046 |
| Laiiguagelog | 1460 | 1004 | 75 | 15.936 | 0.212 |
| Bibtex | 7395 | 1836 | 159 | 2.402 | 0.015 |
| Cal500 | 502 | 68 | 174 | 26.044 | 0.150 |

*(2) Ease of Operation*. Interface design: From the perspective of the user interface, in the online English reading system, the most important factor that affects the end user’s reading experience is the design of the interface itself, and the most important thing is the font used to express the content and color of the reading. For color matching, the system interface design uses a bright dark green background, bright blue fonts, and gray water lines. This color combination gives people a soft feeling and helps students successfully complete reading tasks [17]. The fonts are the most common English Arial and Chinese Song Ti, and the font size is 10.5 pt.

*(3) Scalability*. Function expansion: During system design, interfaces are reserved for each module to facilitate integration with other platforms and systems. For example, the user management in the system management module communicates user data with education management systems in other schools and other internal management information systems, thereby providing a complete interface for unified user identity verification [18].

Software and hardware upgrade: Based on the design of the system, it is assumed that the development and production environments used in the system are all open-source software, including the current very mature server architecture LAMP or LNMP. Functions such as reverse proxy and load balancing are very easy to use and basically guarantee more system carrying capacity in the future.

*(4) Openness*. The system data format complies with relevant national and industry standards, thus ensuring good compatibility and application portability [19]. As mentioned above, the university’s English online reading system uses the B/S architecture to develop the system, so all graphical user interfaces (GUIs) are displayed on the web page in XHTML format.

In order to facilitate maintenance and the separation of business logic and user interface, the interface uses XHTML to organize content, and CSS style sheets are responsible for visually presenting specific pages [20]. Human-computer interaction is mainly realized by using script libraries such as JavaScript or jQuery.

#### *4.1.2. Security Requirements*

*(1) Fault Tolerance*. If incorrect data are generated due to user input or operation errors, the system will automatically send the log to the designated administrator mailbox by e-mail.

*(2) The External Environment Safety of the System*. The design and operating environment of the university’s network reading learning system are based on the university network, which is logically separated from the external network, thereby providing basic network security [21]. The server uses the Ubuntu Linux operating system, which has a built-in software firewall and provides a complete privilege management system at the operating system level.

*(3) Internal Security of the System*. While protecting the external system, the system provides a complete mechanism for user authentication and privilege management so that users who meet the access requirements can obtain appropriate resources to train the system without interference [22]. In this case, users who are not authenticated or whose permissions do not match will be completely blocked in the system.

*(4) System Operation Safety*. According to design assumptions, the server that runs the system is located in the school’s central computer room, and a complete disaster recovery and emergency plan has been developed. UPS uninterruptible power supply and generator set provide power supply, and the server is also equipped with redundant drives.

*4.2. System Function Design.* When analyzing the needs of college English reading comprehension by decomposing the data flow diagram of each level, a hierarchical diagram (H-diagram) of the system’s functional structure can be obtained.

The online English teaching system mainly includes a course and problem management module, a user dictionary module, a system management module, a performance management module, an automatic shopping module, and a teaching module. Among them, system management mainly meets the functional requirements of administrators to import, modify, delete users, maintain, and configure system parameters through the system. Performance management mainly meets the functional needs of managers and teachers in performance monitoring, responsible for reviewing and evaluating objective questions in automated scoring systems [23]. Learning to read mainly meets the following functional requirements: setting educational tasks, teaching business, testing business, creating an educational content interface, and recording the learning process. User dictionary will be added to the user list, unfamiliar English vocabulary, and query requirements; course and question bank management provides a management interface for teachers, allowing teachers to maintain and update learning content and test question banks.

### 4.3. Database Design

*4.3.1. Semester Information.* The basic information of the semester is saved.

The semester ID is the primary key of the semester information table in Table 3. It is a column or columns that uniquely identify each row in the table.

*4.3.2. Course Information.* The basic information of the course is saved.

The course ID is the primary key of the course information table in Table 4. It is a column or columns that uniquely identify each row in the table.

*4.3.3. Unit Information.* The basic information of the unit is saved.

The unit ID is the primary key of the device information table in Table 5. The course identifier is a column that establishes and strengthens the data link between the course information table in Table 4 and the unit information table in Table 5. Therefore, the course identifier is a foreign key in the module information table in Table 5.

*4.3.4. Article Content.* The basic information of the article content is saved.

The article ID is the primary key of the article content table in Table 6. It is a column or columns that uniquely identify each row in the table.

*4.3.5. Learning Record.* The basic information of learning records is saved.

The learning record ID is the primary key of the learning record table in Table 7. It is a column or columns that uniquely identify each row in the table.

*4.4. System Development Environment.* Nonfunctional requirements include the standards, specifications, and contracts that the university's online reading education system must meet, the specific details, and performance requirements of the front-end interface, analysis, and design constraints, and quality attributes. In order to meet the needs of multiple people watching online learning videos at the same time, the Linux operating system is used to ensure the real time and security of the server, and the use of MySQL for the database can save costs and improve the security of data sharing, ensuring that teachers and students use the system in the process. The client can use any operating system without revealing privacy. The most suitable WIN7 operating system is recommended, and the browsers are also compatible. The most suitable browser is Explorer 11, because it is more compatible with classroom video playback and not easy to crash.

The server-side is shown in Table 8:

The client is shown in Table 9:

*4.5. Application Advantages of English Online Teaching Mode.* Teachers are the creators and organizers of educational activities. The most important thing is that all teachers are testers of teaching effectiveness. In the era of internet + education, English classrooms are facing new opportunities and challenges in the classroom. Teachers' computational thinking is essential for the effective use of the internet. What is computer thinking? According to Professor Jeannette M. Wing, Dean of the School of Computer Science at the Carnegie Mellon University in 2018, in the prestigious ACM paper, computational thinking is a basic skill that everyone should have in the digital age. "Until 2050, every citizen on the planet must have the ability to think using computational methods" [24], not the special skills of a computer scientist. Noncomputer English teachers have improved their computational thinking skills, using the educational resources of the online learning platform to teach and publish homework, and use the platform's data processing capabilities to understand students' homework needs. Technology can play a real role [25]. These are all interdisciplinary technologies, and many technical applications and related solutions are not yet perfect. English teachers face bottlenecks when using this education platform. English online education platform, this subject platform, is only an auxiliary tool for English teaching. Teachers need to fully understand the various functions of the platform and other functions, and the English online education system platform can provide synergy [26]. Otherwise, the English online education platform can only be used in internet search engines. If you are not familiar with the functions of the platform, the completion of classwork is restricted. The teachers were always busy clicking and searching in the classroom, and after spending a lot of time, the students fell into chaos [27]. Therefore, this violates the original intention of the employees to design and develop the Xueleyun online learning platform. Therefore, although the platform is powerful, it poses new challenges for English teachers. The workflow of the platform should be mastered, how to use each functional area should be clarified, and they must be familiar with creating e-learning resources [28]. These requirements and the resulting technical applications also make it impossible for English teachers to use the education platform. Technology is a tool but also a double-

Table 3: Semester information form.

| Field name | Type of data | Length | Key | Description |
|---|---|---|---|---|
| Semester ID | Int (2) | 2 | Primary key | Semester ID |
| Semester name | Varchar (40) | 20 | No | Semester name |
| Semester open date | Date | | No | Semester start date |
| Semester close date | Date | | No | Semester end date |

Table 4: Course information form.

| Field name | Type of data | Length | Key | Description |
|---|---|---|---|---|
| Course ID | Int (2) | 2 | Primary key | Course ID |
| Course name | Varchar (50) | 50 | No | Course title |
| Course code | Varchar (30) | 30 | No | Course code |
| Course textbook | Varchar (80) | 80 | No | Course materials |
| Course description | Text | | No | Description |

Table 5: Unit information table.

| Field name | Type of data | Length | Key | Description |
|---|---|---|---|---|
| Unit ID | Int (3) | 3 | Primary key | Unit ID |
| Unit name | Varchar (50) | 50 | No | Unit name |
| Unit title | Varchar (100) | 100 | No | Unit title |
| Course ID | Int (1) | 1 | Foreign key | Course ID |

Table 6: Article content.

| Field name | Type of data | Length | Key | Description |
|---|---|---|---|---|
| Text ID | Int (5) | 5 | Primary key | Article ID |
| Text title | Text | | No | Article title |
| Text content | Text | | No | Article content |
| Text author | Varchar (60) | 60 | No | Author |
| Text url | Varchar (200) | 200 | No | Source |
| Text unit | Int (3) | 3 | No | Belonging to the unit |
| Text word count | Int (5) | 5 | No | Article word count |
| Textbook | Int (2) | 2 | No | Affiliated textbook |

Table 7: Learning record.

| Field name | Type of data | Length | Key | Description |
|---|---|---|---|---|
| Leaming ID | Int (10) | 10 | Primary key | Learning record ID |
| User ID | Int (5) | 5 | Foreign key | Student ID |
| Text ID | Int (4) | 4 | Foreign key | Article TD |
| Leaming start time | Date time | | No | Starting time |
| Leaming end time | Date time | | No | End time |
| Leaming speed | Int (4) | 4 | No | Reading speed |
| Leaming answer rec | Text | | No | Answer content |
| Leaming score | Int (3) | 3 | No | Learning score |
| Leaming count | Int (l) | 1 | No | Number of learning |

Table 8: Server-side environmental parameter table.

| | |
|---|---|
| Database used by the system | MySQL 5.5. 35 |
| Operating system | 64-Bit version of Ubuntu Linux server (64 bit) 12.04.3 |
| Web server | Apache 2.4,9 |
| Database | MySQL 5.5. 35 |

Table 9: Client environmental parameter table.

| | |
|---|---|
| Operating system | Windows 7 Professional SP1 |
| Browser | Internet Explorer 11 |
| Resolution | The best effect is 1024 × 768 pixels |

edged sword. Only teachers can creatively use these technologies, and online learning platforms such as English online education system platforms can make full use of these technologies and improve the quality of English classroom teaching [29]. If you blindly rely on these internet technologies, download materials, analyze materials, and assign homework, then these advanced technical tools will hinder the development of teachers' classroom teaching skills.

The new classroom based on the English online education system platform provides students with language materials of different forms and functions, vivid photos, and interesting music through the combination of photos, text, and audio. Learning has aroused people's strong interest in learning and attracted the attention of students. In the classroom, students use "eyes, ears, mouth, brain, and hands" and combine them with the actions of "seeing, listening, speaking, and thinking" [30]. Through various forms of audiovisual teaching, the educational goals of students' language expression ability can be fully explored.

In short, the English online education system platform is a daily learning platform for teachers and students. It explores the essence of student learning and makes daily learning easier and more effective. It exists in the whole process of classroom teaching, promotes the development of classroom informatization, optimizes and perfects traditional teaching methods. One of the characteristics of the new goal of "Go for it" [31] is to eliminate emotional confusion and ignorance by instilling emotional education in the classroom and allowing students to release their emotions during teaching and discussion.

In the process of researching new education models, the education evaluation system must also appropriately adapt to the requirements of education reform. The immutable educational evaluation no longer adapts to the times. In the process of English teaching reform, it is necessary to flexibly change the evaluation standards, evaluation methods, and evaluation objectives [32]. Moreover, all teachers are constantly growing. It is impossible to use the same set of indicators to evaluate teachers of different levels, which ignores the gradual characteristics of teachers' personal growth. In terms of education evaluation, the UK has an evaluation system for curriculum development, education content and organization, education system and verification, student development and learning outcomes, quality assurance, and improvement mechanisms [33].

Therefore, at the dawn of the internet + education era, all first-class English teachers are advancing with the times, optimizing the learning process, improving the application value of their educational abilities, and actively developing new models. In the process of learning and exploring English classroom teaching in the new era, continuous self-development can become an English teacher with strong adaptability in the process of the new era.

## 5. Conclusions

Multilabel learning is a hot topic in machine learning research, but scientists ignore the inherent imbalance of multilabel datasets and focus more on the relationship between labels. The problem of category imbalance is another hot spot in machine learning research. The most important thing is that imbalanced data will distort the classification aspect of the classification model and reduce its effectiveness. However, this study on categories is mainly focused on the imbalance between the two traditional types. Therefore, it is necessary to urgently solve the problem of imbalance in multilabel classification. This study first discusses the basic knowledge of research and the importance of multicomponent learning with unbalanced categories and then introduces related methods to deal with multicomponent learning with unbalanced categories. Finally, this study explores two different fields of university reading education and software development and uses the latest information technology to improve the learning of university English courses. It tries to analyze a series of needs and use the B/S structure for systematic research and analysis. The system analyzes the functional modules of teaching reading, the functional modules for automatic evaluation, the score management module, the user dictionary, the curriculum and question library management, the system management module, and other modules. This allows each module to achieve the goal of the design function and ultimately achieve the overall interaction. Today, blockchain technology has become an anticounterfeiting and tamper-proof technology. In the future, the CSD-ELM technology used in this study can be combined with blockchain technology to complete the real-name system of online English teaching and ensure the authenticity of students' learning.

## Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

[1] M. Huang, "Factors affecting the effectiveness of online English classroom teaching and suggestions for improvement," *English Teacher*, vol. 21, no. 01, pp. 159–161, 2021.

[2] W. Huang, "Practical exploration of the combination of online teaching and classroom teaching in colleges and universities," *Modern English*, vol. 12, no. 02, pp. 16–18, 2021.

[3] C. Wang, "Discussion on the interactive mode of college English online classroom," *University*, vol. 23, no. 03, pp. 55-56, 2021.

[4] W. Zhou, "Some thoughts on English teaching in junior high schools under the network environment," *Educational Circles*, vol. 26, pp. 12-13, 2021.

[5] L. Wei, "The application of blended teaching mode in basic English courses under the background of the Internet: a review of "English online classroom teaching mode and method research"," *Journal of Tropical Crops*, vol. 42, no. 12, 3757 pages, 2021.

[6] L. Zhang, "Research on the construction of three-dimensional network classroom teaching mode for college English," *Modern English*, vol. 19, pp. 25–27, 2021.

[7] L. Liu, "The reform path of higher vocational English teaching based on network classroom teaching," *Hunan Education (C Version)*, vol. 36, no. 08, pp. 45–47, 2021.

[8] L. Xunfang, "Analysis on the management of network classroom teaching in multimedia classroom teaching," *International Journal of Intelligent Information and Management Science*, vol. 7, no. 3, 2018.

[9] L. Zhong, X. Fan, and Z. Liu, "Micro-interaction research on online classrooms for design majors," *China Light Industry Education*, vol. 24, no. 05, pp. 87–90+96, 2021.

[10] C. Guan, S. Xu, and X. Wang, "Analysis of internet-based course teaching practice methods," *Integrated Circuit Applications*, vol. 38, no. 12, pp. 102-103, 2021.

[11] Y. Zhu, "Construction and practice of urban and rural synchronous online classroom teaching mode," *Science and Education Wenhui (Late Issue)*, vol. 10, pp. 160–163, 2021.

[12] R. Zhu, "Reflections on the teaching reform of commercial law courses under the background of the Internet age," *Education and Teaching Forum*, vol. 40, pp. 66–68, 2021.

[13] H. Wu, "Research on strategies to improve the quality of online mathematics classrooms in junior high schools," *Mathematics Learning and Research*, vol. 26, pp. 40-41, 2021.

[14] T. Weicheng, "Research on the strategy of cultivating students' innovative ability in middle school computer teaching," *Everyday Love Science (Teaching Research)*, vol. 15, no. 09, pp. 5-6, 2021.

[15] C. Fu, "How to establish online classroom teaching of art in primary schools," *Xue Weekly*, vol. 25, pp. 171-172, 2021.

[16] S. Li, C. Gong, W. Dong, and B. He, "Research on online teaching of biochemistry practice courses in colleges and universities during the epidemic period," *Guangzhou Chemical Industry*, vol. 49, no. 15, pp. 246–248, 2021.

[17] X. Meng, C. Hui, P. Huang, and Y. Deng, "Investigation and analysis of online classroom teaching quality of automobile majors in higher vocational colleges and countermeasures," *Auto Interiors*, vol. 15, pp. 32-33, 2021.

[18] W. Li, "Some thoughts on network teaching in primary and secondary schools," *Contemporary Family Education*, vol. 22, pp. 59-60, 2021.

[19] W. Ren, D. Kang, Y. Tang, and A. B. Chan, "Conference on computer vision and pattern recognition," in *Proceedings of the Conference on Computer Vision and Pattern recognition IEEE CVPR*, pp. 5353–5362, Wellington, New Zealand, 2018.

[20] M. Ullah and F. A. Cheikh, "2018 IEEE international conference on image processing," in *Proceedings of the International Conference on Image Processing IEEE ICIP*, pp. 3738–3742, Athens, Greece, October 2018.

[21] M. Ullah and F. Alaya Cheikh, "International conference on computer vision and pattern recognition," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition Workshops IEEE CVPRw*, pp. 1816–1823, Wellington, New Zealand, 2018.

[22] J. Liu, C. Gao, D. Meng, and A. G. Hauptmann, "Conference on Computer vision and pattern recognition," in *Proceedings of the Computer Vision and Pattern recognition IEEE CVPR*, pp. 5197–5206, Wellington, New Zealand, 2018.

[23] H. Ullah, M. Ullah, and M. Uzair, "A hybrid social influence model for pedestrian motion segmentation," *Journal of Neural Computing and Applications*, vol. 29, pp. 1–17, 2018.

[24] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[25] G. Cheng, C. Yang, X. Yao, L. Guo, and J. Han, "When deep learning meets metric learning: remote sensing image scene classification via learning discriminative CNNs," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 5, pp. 2811–2821, 2018.

[26] Z. Zhang, J. Geiger, J. Pohjalainen, A. E. Desoky Mousa, W. Jin, and B. Schuller, "Deep learning for environmentally robust speech recognition: an overview of recent developments," *Transaction on Intelligent Systems and Technology: ACM*, vol. 9, no. 5, p. 49, 2018.

[27] J. Su, "Talking about the application of online classroom in English teaching," *Shandong Education*, vol. 29, pp. 61-62, 2021.

[28] H. Ullah, S. D. Khan, M. Ullah, F. A. Cheikh, and M. Uzair, "2019 8th European workshop on visual information processing," in *Proceedings of the 2019 8th European Workshop on Visual Information Processing (EUVIP)*, pp. 93–98, IEEE, Paris, France, 2019.

[29] J. Shao, C. Change Loy, and X. Wang, "Conference on computer vision and pattern recognition," in *Proceedings of the IEEE 2019 8th European Workshop on Visual Information Processing CVPR*, pp. 2219–2226, Wellington, New Zealand, 2014.

[30] X. Lin, "Research on countermeasures for improving teachers' online classroom management ability," *Xueyuan*, vol. 14, no. 19, pp. 48–50, 2021.

[31] X. Zhang, Q. Zhang, S. Hu, C. Guo, and H. Yu, "Energy level-based abnormal crowd behavior detection," *Sensors*, vol. 18, no. 2, p. 423, 2018.

[32] H. Duan, S. Qu, and X. Zhang, "Research on the development of classroom networking in higher education," *Heilongjiang Science*, vol. 12, no. 13, pp. 29–31, 2021.

[33] H. Wang, "Explore how to improve the teaching effect under the network teaching mode," *Anhui Education and Research*, vol. 18, pp. 95-96, 2021.