



## Research and Design of College English Online Teaching based on Computer Aided Instruction Resource Library

Caifang Huang<sup>1,\*</sup>  and Shuzhen Wu<sup>2</sup> 

<sup>1</sup>Department of Foreign Languages, Hubei Engineering University, Hubei 432000, China, [hcf5200@hbeu.edu.cn](mailto:hcf5200@hbeu.edu.cn)

<sup>2</sup>Department of Computer Science, Hubei Engineering University, Hubei 432000, China, [wsz197906@163.com](mailto:wsz197906@163.com)

Corresponding author: Caifang Huang, [hcf5200@hbeu.edu.cn](mailto:hcf5200@hbeu.edu.cn)

**Abstract.** Online education colleges or English training institutions begin to develop English courses for different devices, which can be used for learners to learn English. In these platforms of English Language Teaching (ELT), all kinds of data are produced, such as learners' grades, learning records, communication records and so on. Although the progress of database technology makes it easier to collect and store data, the explosive growth of data scale is far beyond people's comprehension, and the hidden information in these data can't be found by using traditional database administration system and previous analysis methods. In this article, a data mining (DM) algorithm of English online education platform based on deep learning (DL) is proposed. With more targeted learning plans and instructional resource recommendations. The results show that the recommendation accuracy of this algorithm is over 95% after repeated iterations. DL-based educational DM model can obtain meaningful information from massive teaching data and provide a more scientific decision-making basis for ELT managers.

**Keywords:** English Teaching; Online Education; Data Mining; Deep Learning.

**DOI:** <https://doi.org/10.14733/cadaps.2023.S7.175-185>

### 1 INTRODUCTION

With the deepening of IT in ELT field, a large number of digital ELT resources make learners face difficulties in choosing and using them. At the same time, in the practical application and promotion of English online instructional resources, many disadvantages in design and application are gradually exposed [1]. In order to manage these data, large databases are widely used in commercial and scientific engineering fields. Although the progress of database technology makes it easier to collect and store data, the explosive growth of data scale is far beyond people's comprehension, and the hidden information in these data can't be found by using traditional database administration systems and previous analysis methods [2]. The existing network-assisted instructional resources attach importance to the transmission of English knowledge, emphasize the

input of learning forms, and lack reasonable teaching design and interactive teaching experience. Learners passively and mechanically accept the learning content and memorize or repeat it, which makes it difficult to achieve deep processing, deep understanding, long-term maintenance and flexible application. Emerging DM can discover some unknown and valuable rules from massive data, which undoubtedly provides strong support for providing personalized teaching services [3]. Combining database, data warehouse and other technologies, using DM, we will make full use of learners' academic records, browsing patterns, online records and other data to obtain learners' learning characteristics, and turn educators' teaching experience into strategies and rules that computers can operate.

Under the background of the great times, there are many ways to learn English, and there are many software that can be used for English learning. Online education colleges or English training institutions have started to develop English courses for different devices, which can be used for learners to learn English. In these English learning platforms, various data, such as learners' grades, learning records, communication records, etc., have been produced, many of which have been shelved, and educators or teaching administrators have not fully utilized them [4]. The data sources of DM are diverse, and they must be real and abundant. The result of DM is to acquire knowledge, which is based on the data source, analyzing and reasoning the data, so as to acquire discover the knowledge that users are interested in [5]. This article puts forward the DM algorithm of English online education platform based on DL. By mining the data of learners' learning behavior, it analyzes learners' learning habits and preferences, more targeted learning plans and instructional resource recommendations.

With the increasing popularity of English on-line instruction, a large amount of on-line instruction data has been piled up, so it is very necessary to apply DM to on-line instruction. DM can help us find the problems existing in on-line instruction from the perspective of data, and objectively reflect the problems in on-line instruction platform. To improve the quality of on-line instruction, DM can find unknown knowledge, which is different from those data processing methods that first put forward assumptions and then verify them. The validity of mining results means that the DM results are correct and reasonable, and the operability of mining results means that the mined knowledge can provide support for decision-making. The main innovations of this article are as follows:

(1) Starting with the value and significance of educational big data, this article discusses the strategy of constructing the DM model of English online education platform.

(2) Through the practice of mining in the teaching model, the research has obtained valuable information, thus providing learners with more targeted learning plans and instructional resource recommendations.

## 2 RELATED WORK

Huang et al. applied particle swarm optimization algorithm to personalized recommendation of technical articles in the Internet environment, and recommended learning resources such as technical articles in the Internet environment to learners by establishing learning objectives, subscription information and other attributes, matching with the number of comments and difficulty of articles. Countries all over the world have strengthened the development of education to promote the information process. Computer assistance can make up for the lack of specialized teachers in English teaching for students in China. Standard speech input can be accepted in both urban and rural areas to improve students' ability of speech imitation. Simulation technology can bring students a sense of immersive, so that they can carry out various simulation practice activities in a safe environment. Li [6] has developed an English learning system. The language meaning is constructed through more meaningful information to enhance the understanding of semantics. This can not only improve students' interest in English in the teaching process. You can also use the interactive functions of various software to provide students with English situations for independent exploration, or you can design a variety of multimedia courseware for students to

click and learn. Ma and Ding [7] Computer aided English education believes that education consists of various complex interactive relationships. In addition, computer aided learning also builds a learning environment to encourage students to learn independently. Only with the participation of various factors in the environment can the interactive relationship be realized. Therefore, he attaches great importance to the creation of children's creative learning environment. Therefore, the use of multimedia computer assisted English teaching is the main way to reform traditional English teaching. Shi [8] If teaching designers and kindergarten teachers take these factors into consideration, they should choose appropriate information presentation and student activities according to the characteristics of English teaching development. It is necessary and of great practical significance to make use of computer multimedia to create an English learning environment so as to stimulate interest in English learning and improve English teaching. It is necessary to avoid the negative impact of technology and make it fully used by us, which will greatly promote children's English learning. Shu [9] thinks that the learning environment should be divided into classroom basic learning environment, computer micro environment and open virtual learning environment. It can also be called an e-learning environment, which almost covers all our current understanding of the learning environment. The open virtual learning environment can be understood as the opening of teaching concepts. Diversification of learning styles. Therefore, students should take the open teaching concept as the guidance and use the advantages of computer multimedia to create a more vivid teaching situation in the basic classroom environment. Computer micro environment is a learning environment supported by modern information technology and abundant network resources. Wang and Zhao [10] think that learners have different ways to obtain learning information or learning resources. The environment and learning objectives are constantly changing. Therefore, our teaching design should also be flexible and open. Knowledge should choose the questions related to the learners' life experience. At the same time, provide tools to better understand and solve problems, and guide them to take the initiative. this algorithm is applied to analyze learners' test scores, so that learners' scores can be clustered according to a certain number, and the results of clustering analysis can be used to guide learners' learning and future teaching work. Chen et al. [11] established the performance evaluation index system of colleges and universities by using cluster mining method, and on this basis, divided the schools with similar evaluation indexes into one class, and built a favorable decision support environment for the performance evaluation of colleges and universities.

This article analyzes the importance of educational big data in intelligent education, explores the application of educational big DM in ELT management, and establishes learner portrait model and learning resource portrait model combined with DL algorithm. The model builds a model for individual learners, divides learning resources into knowledge points, and provides individualized teaching plans for learners, thus realizing the construction of online education platform for smart English.

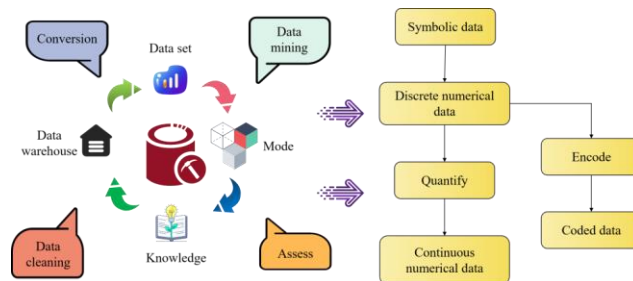
### **3 METHODOLOGY**

#### **3.1 Application of DM in English Online Education Platform**

English learning is a step-by-step process, and the learning content corresponding to each level has a sequential relationship. In the stage of learning, we must advance step by step. Otherwise, it will lead to disjointed learning with little effect. There is a huge contradiction between the difference of educational personality characteristics and the singleness of traditional distance education system services. The traditional distance education platform is not enough to fundamentally solve this contradiction. In order to improve the utilization of educational information resources, fully meet the teaching needs and provide personalized teaching services for learners. Some learners' knowledge of English should be at a low level, but they don't have a clear understanding of themselves, and they choose advanced courses without a certain ability assessment. Obviously, they are struggling in learning, which will not only take a lot of

unnecessary time, but also make learners slack off, which is not conducive to the sustainable learning of English.

The learning contents in the online English learning platform correspond to specific English grades, and the learning contents corresponding to different English grades include multiple units. The learning content data corresponding to each grade is very huge, so the association analysis data only extracts the learning content of one unit in a certain registration from the database. From the epistemological point of view, knowledge is a general description of things' moving state and its changing law. However, based on artificial intelligence and information systems, this definition needs to be expressed more accurately. Knowledge comes from the classification ability of human beings. Knowledge about the environment is the complex classification of sensory signals from the viewpoint of survival. The data representation and transformation process in DM is shown in Figure 1.



**Figure 1:** Data representation and transformation process in DM.

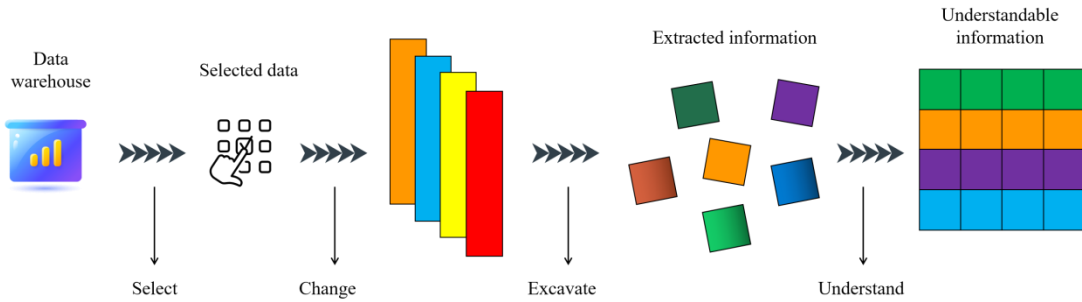
The specific application of DM in personalized distance education system makes the resources of distance education system be allocated around the needs of learners, and supports the individualized learning needs of learners. The personalized distance education system is an intelligent system, which can learn more about learners' learning characteristics, ability differences, hobbies and so on, and dynamically adjust teaching strategies according to these factors, so as to truly teach learners in accordance with their aptitude and realize personalized teaching.

### 3.2 DM Algorithm for Online English Education

English learning is a step-by-step process, and each level of learning has a sequential relationship. For on-line instruction adults, they only have a general understanding of their English level and don't know which level they are in, which requires educators to guide them in learning. However, the existing network teaching staff can only give learners some subjective guidance based on their knowledge or experience, but there is no more accurate quantitative data to guide learners on how to make them learn English step by step. Nowadays, many online English learning systems in online education colleges or English learning institutions in the society are that learners choose their own courses and can learn by paying a fee. In the online English learning system, most of the courses are presented by doing various exercises, so that a large amount of achievement data will be accumulated in the learning stage of learners. These data are stored in the database, but they are not used effectively. Teachers or teaching administrators only get information on the surface of many data through simple statistics or sorting, which may be due to the lack of information awareness or IT, resulting in the knowledge hidden in these massive data can not be fully and effectively utilized.

Neural network has many excellent properties. Its plasticity, adaptability and self-organization make it have strong learning ability. Its parallel processing mechanism makes it a short time to solve problems, and it has the potential to meet the real-time requirements. Its distributed storage

mode makes its robustness and fault tolerance quite good. The environment and steps of DM based on DNN are shown in Figure 2.



**Figure 2:** Environment and steps of DM based on DL.

The model uses the  $P$ -dimensional max-pooling operation to aggregate the local important information in the convolutional hidden layer sequence  $h^c$ , thereby obtaining a global-level hidden layer sequence:

$$h^{cp} = \{h_1^{cp}, \dots, h_{\lfloor (N+k-1)/p \rfloor}^{cp}\} \quad (1)$$

Among them, each hidden layer element is:

$$h_i^{cp} = \left[ \max \begin{bmatrix} h_{i-p+1,1}^c \\ \dots \\ h_{i,1}^c \end{bmatrix}, \dots, \max \begin{bmatrix} h_{i-p+1,d}^c \\ \dots \\ h_{i,d}^c \end{bmatrix} \right] \quad (2)$$

E-learning is a kind of learning with many interactions. In terms of interaction objects, who have completed autonomous learning, the discussion of learning methods, the exchange of learning experiences, or the joint efforts to complete an academic task, and the realization of collaborative learning among learners. Constructing characteristic difference evaluation function according to recommendation model:

$$\min F(x) = \sum_{i=1}^3 W_i F_i \quad (3)$$

$$F_1 = \frac{\sum_{n=1}^n x_{mn} |S_{mn} - DI_{mn}|}{\sum_{n=1}^N x_{mn}} \quad (4)$$

$$F_2 = \frac{\sum_{k=1}^k |MF_{mk} - M_{mk}|}{k} \quad (5)$$

$$M_{mk} = \frac{Num2_{mk}}{\sum_{k=2}^k Num2_{mk}} \quad (6)$$

$$Q2 = X_{mn} MT_{mn} \quad (7)$$

$$F_3 = \frac{\sum_{k=1}^k |CF_{mk} - C_{mk}|}{k} \quad (8)$$

$$C_{mk} = \frac{Num3_{mk}}{\sum_{k=1}^k Num3_{mk}} \quad (9)$$

$$Q3_{mn} = X_{mn} CT_{mn} \quad (10)$$

Because each learner is different in knowledge level, cognitive ability, learning style and learning motivation, the learning strategies adopted in the learning process are also different. However, learners are not clear about which learning strategies to adopt because of their lack of self-control learning ability, so they need guidance and help on how to choose learning strategies.

Use  $u_{ij}$  to represent the results of the English classroom comparison. After comparing all elements at each level pairwise, a pairwise comparison judgment matrix is obtained, and the matrix is expressed as follows:

$$U = (u_{ij})_{n \times n} = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \dots & \dots & \dots & \dots \\ u_{n1} & u_{n2} & \dots & u_{nn} \end{bmatrix} \quad (11)$$

Calculate the weight of each sub-factor, and verify the consistency by solving the maximum eigenvalue of the judgment matrix. First, calculate the normalization for each column:

$$\bar{u}_{ij} = \frac{u_{ij}}{\sum_{k=1}^n u_{kj}} \quad (12)$$

Second, average the normative columns to determine the final weights:

$$\hat{w} = \frac{1}{n} \sum_{j=1}^n \bar{u}_{ij} \quad (13)$$

The feature vector is the weight of each factor:

$$\hat{w} = (\hat{w}_1, \hat{w}_2, \dots, \hat{w}_n) \quad (14)$$

Compute the consistency metric for the constructed matrix:

$$CI = \frac{\lambda_{\max} - n}{n-1} \quad (15)$$

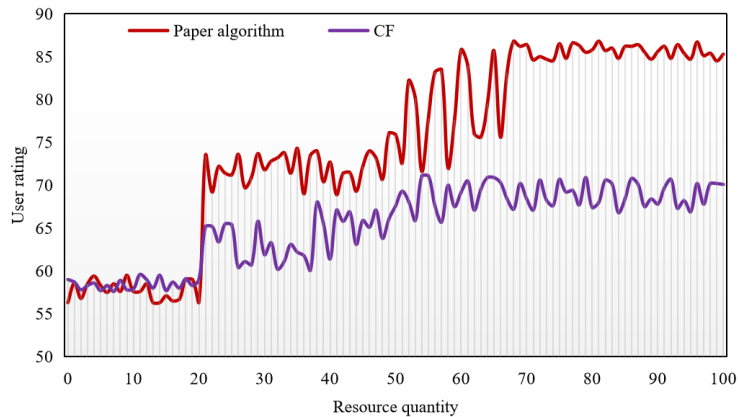
Compute the largest eigenvalue of the judgment matrix:

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^n \frac{(U\hat{W})_i}{\hat{W}_i} \quad (16)$$

$(U\hat{W})_i$  indicates that the  $W$ -th element of vector  $i$  is used.

#### 4 RESULT ANALYSIS AND DISCUSSION

With the support of the network environment, educators and learners can exchange information at any time and any place, without having to stick to the limited time in class. Figure 3 shows learners' subjective grading results of traditional online teaching methods and personalized teaching platforms.



**Figure 3:** Learners' subjective rating.

Most learners said that the personalized teaching platform based on DL improvement can help them quickly locate the information they need in the massive resources, meet their individual needs, and tap learners' potential interests at the same time.

The quality of data is the key to DM, and the quality of mining model construction is determined by the quality of data. Data preparation includes data collection, data mining database establishment and data preprocessing, which refers to the necessary conversion, cleaning, filling and merging of the determined basic data, covering all activities from preliminary rough data to the construction of the final data set, including the selection of table and record attributes, and data trimming and conversion according to the requirements of DM model building tools. Taking ELT resource recommendation accuracy as the test index, CF algorithm and FCA algorithm are selected as the comparison objects. The experimental results are shown in Table 1, Table 2 and Table 3.

<i>Sample size</i>	<i>Resource recommendation accuracy (%)</i>
15	97.68
30	96.88
45	94.38
60	93.21
75	91.95
90	91.23
105	90.74

**Table 1:** ELT resource recommendation accuracy of this method.

<i>Sample size</i>	<i>Resource recommendation accuracy (%)</i>
15	90.02
30	85.75
45	83.5
60	75.38
75	71.13
90	69.85
105	67.86

**Table 2:** ELT resource recommendation accuracy of CF algorithm.

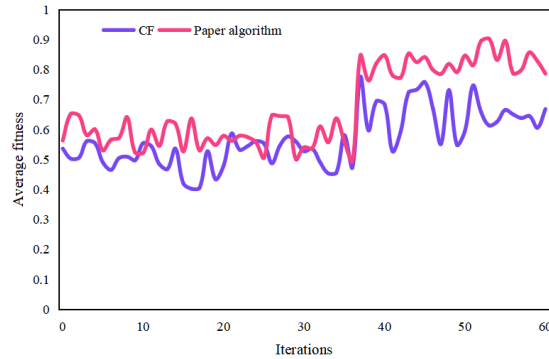
<i>Sample size</i>	<i>Resource recommendation accuracy (%)</i>
15	88.55
30	87.53
45	83.75
60	76.88
75	74.31
90	70.01
105	66.85

**Table 3:** ELT resource recommendation accuracy of ID3 algorithm.

When the number of test samples began to increase, the resource recommendation accuracy of different DM methods showed a downward trend. However, compared with the traditional CF and ID3, the instructional resource recommendation accuracy of this method is obviously higher. After repeated iterations, the recommendation accuracy of our algorithm is over 95%. Therefore, it is of positive significance to apply this model to the teaching management of English online education platform.

No matter the video of guiding learners before class or the projects and cases presented in class, it is necessary to create real life situations and meaningful English learning situations for learners, that is, to embed learners' learning into real and complex problem situations, to realize the situational, problematic and task-oriented learning, to effectively promote the construction of knowledge in real situations, and to transfer, apply and create in similar situations. After network training, a fully connected network that can correctly classify input tuples is obtained. But such networks usually have too many neurons and connections. Even for a network with only one output neuron, with the increase of input neurons, the number of connections between neurons in each layer of the network will increase exponentially, resulting in the exponential growth of the corresponding extraction rules, which makes it very difficult to extract rules. Therefore, the network must be cut before extracting neural rules. The comparison results of recall rates of different algorithms are shown in Figure 4.

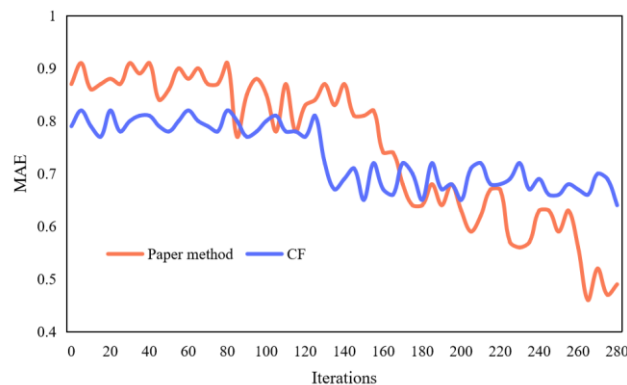




**Figure 4:** Comparison results of recall rate.

The results show that this algorithm can obtain more reasonable, feasible and scientific teaching management effect than the standard traditional DM algorithm. The algorithm in this article is used to realize English personalized teaching, which has good optimization characteristics and fast convergence speed.

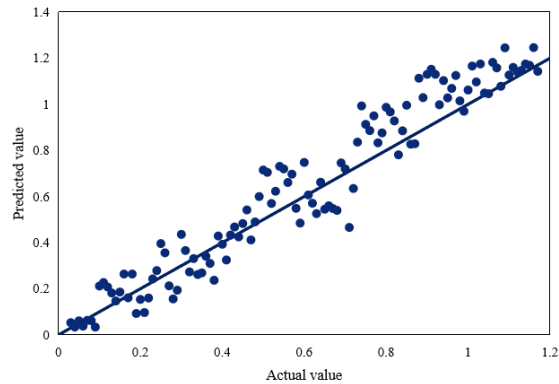
Most databases are relational, so it is very important to effectively execute DM on relational databases. But there are different data and databases in different fields, and they often include complex data types. DM points out that it discovers new knowledge and can express it in the form of relations, rules, concepts, etc., but users don't know the basic principles of knowledge discovery. Only when DM system can provide a better explanation mechanism can users assess this knowledge more effectively, and distinguish which are really useful knowledge, which are common knowledge or abnormal knowledge. The MAE comparison of different algorithms is shown in Figure 5.



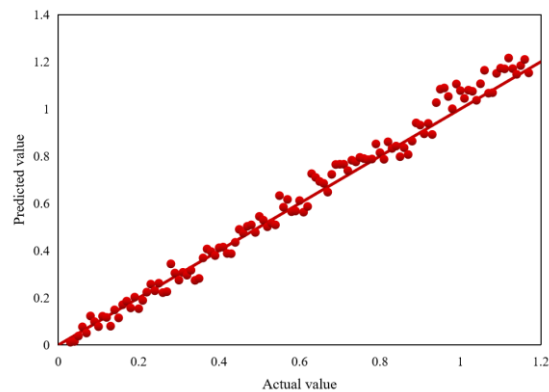
**Figure 5:** Comparison of algorithm MAE.

In the middle and late stage of operation, and the error is significantly reduced. In the stage of establishing the model, you can choose and apply different DM tools and technical methods, compare and assess different types of models, and make a good choice of DM models. For the same DM problem, it is not necessary to stick to a single DM, but to analyze various mining technologies, and we should choose technologies and methods aiming at different data forms for

mining. The test results using the traditional CF algorithm are shown in Figure 6. The test results using the algorithm in this article are shown in Figure 7.



**Figure 6:** Scatter chart of traditional CF actual value and predicted value.



**Figure 7:** Scatter diagram of actual value and predicted value of this algorithm.

DM algorithm of DL-based English online education platform is better than traditional CF in both accuracy and efficiency. The design of digital ELT resources based on DL aims at constructing high-quality digital ELT resources and improving English learners' application level of modern IT and autonomous learning ability. When preparing for DM, it is necessary to assess the existing data, transform the problem to be solved by DM into a DM problem definition, determine the forecast target and make a DM plan.

## 5 CONCLUSIONS

With the increasing popularity of English on-line instruction, a large amount of on-line instruction data has been piled up, so it is very necessary to apply DM to on-line instruction. This article puts forward the DM algorithm of English online education platform based on DL. By mining the data of learners' learning behavior, it analyzes learners' learning habits and preferences, so as to provide learners with more targeted learning plans and instructional resource recommendations. After repeated iterations, the recommendation accuracy of this algorithm is over 95%. When preparing

for DM, it is necessary to assess the existing data and transform the problem to be solved by DM into a DM problem definition. Combining database, data warehouse and other technologies, using DM, we will make full use of learners' academic records, browsing patterns, online records and other data to obtain learners' learning characteristics, and turn educators' teaching experience into strategies and rules that computers can operate.

## 6 ACKNOWLEDGEMENTS

This work was supported by Teaching Reform Research Project of Hubei Engineering University: Research on Teacher-Student Relationship in College English Teaching of application-oriented undergraduate university based on Production-oriented Method (No. 202203); University-Industry Collaborative Education Program of Ministry of Education: The practice of blended College English Teaching Reform in the context of Internet (No. 220503284231550).

Caifang Huang, <https://orcid.org/0000-0003-1737-5952>

Shuzhen Wu, <https://orcid.org/0000-0002-1469-0853>

## REFERENCES

- [1] Agarwal, C.; Chakraborty, P.-A.: Review of tools and techniques for computer aided pronunciation training (CAPT) in English, *Education and Information Technologies*, 24(6), 2019, 3731-3743. <https://doi.org/10.1007/s10639-019-09955-7>
- [2] Gomedede, E.; Gaffo, F.-H.; GU, B.: Application of Computational Intelligence to Improve Education in Smart Cities, *Sensors*, 18(1), 2018, 328-334. <https://doi.org/10.1021/jf051857o>
- [3] Gu, S.; Li, X.: Optimization of computer-aided English translation teaching based on network teaching platform, *Computer-Aided Design and Applications*, 19(S1), 2021, 151-160. <https://doi.org/10.14733/cadaps.2022.s1.151-160>
- [4] Hai, Y.: Computer-aided teaching mode of oral English intelligent learning based on speech recognition and network assistance, *Journal of Intelligent & Fuzzy Systems*, 39(4), 2020, 5749-5760. <https://doi.org/10.3233/JIFS-189052>
- [5] Jing, D.; Jiang, X.: Optimization of computer-aided English teaching system realized by VB software, *Computer-Aided Design and Applications*, 19(S1), 2021, 139-150. <https://doi.org/10.14733/cadaps.2022.S1.139-150>
- [6] Li, H.: Improved fuzzy - assisted hierarchical neural network system for design of computer - aided English teaching system, *Computational Intelligence*, 37(3), 2021, 1199-1216. <https://doi.org/10.1111/coin.12362>
- [7] Ma, H.; Ding, A.: Construction and implementation of a college talent cultivation system under deep learning and data mining algorithms, *The Journal of Supercomputing*, 78(4), 2021, 5681-5696. <https://doi.org/10.1007/s11227-021-04036-4>
- [8] Shi, L.: Application of big data language recognition technology and GPU parallel computing in English teaching visualization system, *International Journal of Speech Technology*, 25(3), 2021, 667-677. <https://doi.org/10.1007/s10772-021-09904-1>
- [9] Shu, Y.: Experimental data analysis of college English teaching based on computer multimedia technology, *Computer-Aided Design and Applications*, 17(S2), 2020, 46-56. <https://doi.org/10.14733/cadaps.2020.S2.46-56>
- [10] Wang, Y.; Zhao, P.: A probe into spoken English recognition in English education based on computer-aided comprehensive analysis, *International Journal of Emerging Technologies in Learning (IJET)*, 15(3), 2020, 223-233. <https://doi.org/10.3991/ijet.v15i03.12937>
- [11] Xu, Y.: Computer-Aided Design of Personalized Recommendation in Teaching System, *Computer-Aided Design and Applications*, 17(1), 2019, 44-56. <https://doi.org/10.14733/cadaps.2020.S1.44-56>