WILEY | Hindawi

Research Article

Analysis on the Development Strategy of Computer Education Microcourse Resources Based on a Cognitive Diagnosis Method

Zhe Yang [1,2]

[1]*School of Computer Science and Technology, Soochow University, Suzhou Jiangsu 215006, China*
[2]*Provincial Key Laboratory for Computer Information Processing Technology, Suzhou Jiangsu 215006, China*

Correspondence should be addressed to Zhe Yang; yangzhe@suda.edu.cn

With the continuous deepening of the new curriculum reform, in order to meet the needs of high-efficiency and high-quality teaching, colleges and universities have carried out continuous reforms and innovations in teaching forms and methods. Microcourse is one of the main innovative forms of curriculum development in colleges at this stage, and it has a great role in promoting the quality of courses and the realization of teaching goals. At the same time, as a new discipline, computer science has its own unique properties. In addition, the traditional cognitive diagnosis model ignores the connection between knowledge points in most cases, and the use of this information is insufficient, resulting in that the traditional model cannot be reasonably applied in some practical teaching scenarios. Under the background of the continuous development of modern curriculum and the continuous innovation of educational methods, cognitive diagnosis, as an important research in educational data mining, plays a pivotal role in student evaluation. Based on this, this paper mainly takes the teaching of computer courses as an example. While introducing the current situation of teaching activities, it studies the specific strategies of microlectures in computer courses, so as to better utilize the advantages of microlectures.

1. Introduction

With the development of the times, people's pace of life is gradually accelerating, and many products of the times that can effectively use fragmented time are born, such as short videos, micronovels, and microblogs. Their common feature is to compress the reading or browsing time to a certain extent, so as to achieve the purpose of simplicity and efficiency. Influenced by such products, a series of "micro" things came into being. Microlecture is a new type of classroom teaching mode derived from the rapid development of information technology [1]. As the name suggests, microlecture is to compress the original lengthy lecture to less than 10-15 minutes. In order to achieve this goal, most speakers will make a prepared speech on a specific topic with the help of multimedia. Due to its short time, real content, new ideas, and other characteristics, microlectures are very consistent with today's fast-paced living habits and have been valued by more and more organizations and individuals, enterprises, and even government agencies. Fragmentation and short time are the two major characteristics of microlecture teaching. It is completely different from the traditional teaching mode, but the teaching effect of the harvest is very obvious. At present, microlectures have been widely used in the teaching of vschools, but in terms of practical application effects, they are not ideal [2]. This paper mainly takes the teaching of computer courses as an example. While introducing the current situation of teaching activities, it also studies the specific strategies of microlectures in computer courses in colleges, so as to better utilize the advantages of microlectures.

Under the background of the continuous deepening of the concept of quality education, the teaching goal of computer courses is not only to let students master the related technologies and applications of computer systems but also to realize students' learning of computer knowledge in a real sense on this basis [3–5]. However, as far as the current situation of computer courses is concerned, due to the deep influence of traditional educational concepts, most schools still follow the traditional full-class teaching mode [6]. The

classroom atmosphere is boring, and it is difficult to motivate students to participate [7]. The desire for classroom activities, coupled with the fact that computer is a highly operable course, relies solely on teachers to explain theoretical knowledge without allowing students to perform practical operations; it is difficult for students to truly grasp the content of what they have learned. Over time, students will lose enthusiasm for learning computer courses [8, 9].

In the teaching of computer courses, most of the questions are subjective questions. Students' thinking and answers to these questions can exercise their computer thinking and ability [10–12]. Therefore, teaching focuses on the process of using what you have learned to solve difficult problems [13]. In the new era, computer teaching in schools should pay more attention to practical teaching, which is convenient for students to master computer operation, and can also greatly improve students' learning enthusiasm and interest. Computer teaching is different from other disciplines. Practice is the fundamental factor for the improvement of computer level. Students should operate more, ask more questions, and understand more in learning, which can promote the improvement of students' computer level and achieve good teaching results. In the field of computer science, there may be only one answer to the problem, but the method to solve the problem, namely, the algorithm, is endless. Therefore, when exercising and inspecting students, it is necessary to write subjective questions with specific ideas, which can better reflect the specific abilities of students [14]. Furthermore, unlike the subjective questions that only give an idea or a solution, the significance of the experiment lies in the specific realization of the function of the algorithm [15]. When answering a question, we only write the rough framework, but in the experiment, we need to implement it. In computer science, completing an experiment means that we have to write the corresponding code and perform the successful operation of the code, in addition to considering the solution process; we also need to pay attention to the syntax of different languages, data interaction, implementation details, and so on [16].

Figure 1 shows the difference between the two methods, with a single numerical value given by the traditional method on the left and a more specific analysis of students' multidimensional abilities on the right [17].

Although cognitive diagnosis has achieved good results in the assessment of students in traditional subjects, its performance in teaching computer courses is not satisfactory [18]. Therefore, this paper mainly takes the microclass as the core and discusses the current situation and specific strategies of computer microclass teaching in colleges, in order to provide a certain reference for the development of this field in the future [19].

2. Related Concepts and Theoretical Basis

2.1. Computer Microcourse Education.
In the process of implementing quality education, microlectures have attracted much attention from all walks of life [20]. Taking the application of multimedia technology as the starting point and core, it truly realizes the close integration between
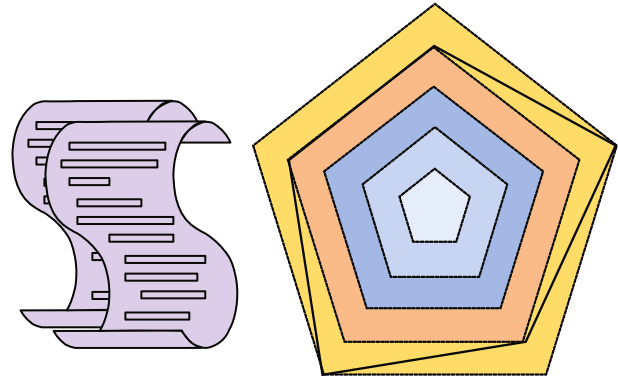

FIGURE 1: Comparison of feedback provided by traditional teaching and cognitive diagnosis.

theoretical teaching and practical teaching. Video plays an important role in microlecture teaching. Academia and theoretical circles have pointed out when researching and analyzing microlectures that microlectures have many characteristics in the process of practical application, which can effectively promote the integration of teaching efficiency and quality.

(1) In the mobile Internet environment, the application of microlectures is becoming more and more extensive, and the results are getting better and better. Compared with other course teaching modes, the video time in microlecture teaching is relatively short, and most of them are short videos. The specific duration of the video is effectively controlled in strict accordance with the cognitive characteristics of students and the actual situation of professional course learning. It is generally controlled within 5 to 10 minutes to better attract students' attention and try to avoid students' lack of concentration in the classroom

(2) The learning tasks involved in microlecture teaching are independent to a certain extent. Each course can be divided into different knowledge units, and teachers can further divide them into different tasks and knowledge points. Video has a certain degree of independence and integrity and can provide targeted tutoring and teaching based on the actual situation of students and provide personalized guidance for students of different levels

(3) The teaching efficiency of microlectures is relatively high and requires the assistance of various mobile devices, which provide students with more opportunities and time for independent learning. Students can make full use of the scattered time to improve their personal academic performance after completing the main learning tasks. In the Internet age, students can watch microclass videos anytime and anywhere and have direct online communication and interaction with teachers based on various

difficulties and obstacles encountered by individuals in the learning process, so as to realize the effective connection between the first classroom and the second classroom and enrich students' extracurricular life. Compared with the traditional mode of lectures, the rich innovative forms of microlectures are also one of the advantages. Speakers can use multimedia, choose to interact, and give a performance style speech. They can also talk for ten minutes like traditional academic reports. For different occasions and different contents, speakers can adopt a variety of forms. They do not have to stick to the limitation of "speakers." What they need to consider is to express their information most efficiently in a limited time

As a creative teaching mode and teaching concept, microcourse meets the actual requirements of national quality education and new curriculum reform, attracts students' attention at the first time, and better realizes the modernization of classroom education. Compared with other disciplines, computer teaching in higher education has higher requirements on students' practical ability and logical thinking ability. Teachers face greater teaching pressure and burden. At the same time, students have many difficulties and obstacles in the process of learning professional courses. If the effective integration between microlecture teaching and higher education computer teaching can be realized, then it can truly break through the shortcomings, improve students' learning efficiency, and create a good environment for students' personal growth.

2.2. Basic Concepts of Cognitive Diagnosis. Cognitive diagnosis theory is a kind of confirmatory model that combines cognitive theory with psychometric model to diagnose the strengths and weaknesses of participants. In short, it is similar to transforming the knowledge points, skills, and thinking process involved in the topic into a matrix, which is set in the model to produce a report on students' mastery of various items. Principle cognitive theory is actually the theoretical basis of the knowledge, skills, thinking, and processing process required by the subjects to answer the right questions. The most important ones are attribute and Q matrix (attribute topic relationship). Cognitive diagnosis is the result of the combination of cognitive psychology and modern measurement, which involves many basic concepts, such as cognitive diagnosis and cognitive diagnosis theory.

The cognitive diagnostic model (CDM) is a psychometric model with cognitive diagnostic function developed. The machining process is measured. Cognitive diagnostic models are constantly being developed. Currently, cognitive diagnostic models can be divided into simplified CDMs and saturated CDMs, parametric CDMs and nonparametric CDMs, and connected CDMs and discrete CDMs. Whether the subject's attribute grasp probability is parameterized or not is the basis for the division of parameterized CDMs and nonparametric CDMs. Among them, there are many and widely parameterized models, such as DINA, RRUM, NIDA, and so on. RSM and AHM are nonparametric

models. The division of simplified CDMs and saturated CDMs is based on the breadth of applicable areas. Simplified CDMs have more constraints or assumptions when constructing project response functions. Models such as DINA, DINO, and NIDA are simplified CDMs; for saturated CDMs, there are GDM, LCDM, and G-DINA. Simplified CDMs are divided into connected CDMs and discrete CDMs according to "how attributes affect item response probability"; corresponding to noncompensated CDMs and compensated CDMs, DINA, NIDA, etc. are connected CDMs and DINO is a discrete CDM.

Deterministic input, noise, and gate model (DINA) is a completely noncompensatory cognitive diagnostic model. It means that the subjects can answer the item correctly only if they have fully mastered all the attributes of the item measurement. The response probabilities of subjects in different categories are shown in Figure 2.

Therefore, the DINA model can only distinguish two types of subjects: one is the fully mastered group and the other is the unmastered group (at least one attribute has not been mastered).

3. Related Technologies

3.1. Probabilistic Graphical Model. A probability graph model is a theory that uses graphs to represent the probability dependence of variables. Combined with the knowledge of probability theory and graph theory, graphs are used to represent the joint probability distribution of variables related to the model. Developed by the Turing prize winner Pearl, probability graph model theory is divided into probability graph model representation theory, probability graph model reasoning theory, and probability graph model learning theory. In recent 10 years, it has become a research hotspot of uncertainty reasoning and has broad application prospects in the fields of artificial intelligence, machine learning, and computer vision. The probabilistic graphical model combines graph theory and probability theory, which uses graphs to represent the interdependent relationships between variables. The probabilistic graph model can visualize the probability model and simplify the mathematical expression of probability. The probability graph contains nodes and edges, where nodes are used to represent one or a group of random variables and edges are used to connect nodes and represent the relationship between nodes' probability relationship. If the edge of the probability graph has a specific direction, we call this kind of probability graph a directed graph model, also known as a Bayesian network; if the edge in the probability graph does not have a specific direction, then this kind of probability graph is called no direct graph model, also known as the Markov random field. We mainly introduce the Bayesian network used in this article.

Considering a joint probability distribution $p(a, b, c)$ over three variables a, b, c, by the product rule of probabilities, we can get

$$p(a, b, c) = p(c|a, b) \cdot p(a, b) = p(c|a, b) \cdot p(b|a) \cdot p(a). \quad (1)$$
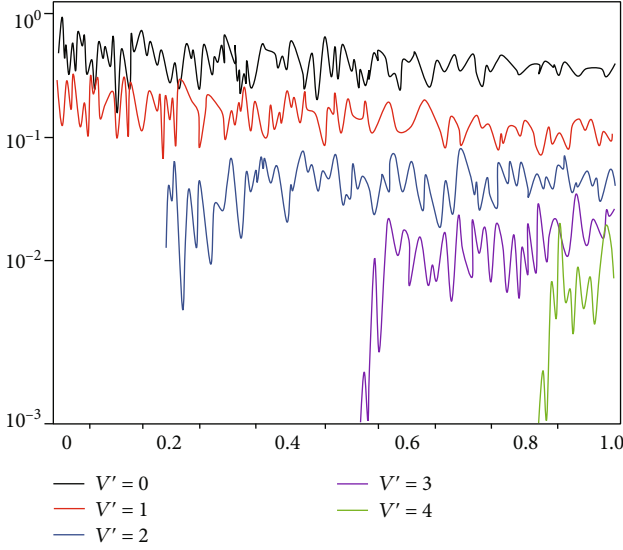
FIGURE 2: DINA model item response probability diagram.

If we change the order of decomposition, we get a different graphical model. Extending the above example, the general relationship between variables in a directed graph and their corresponding probability distributions can be defined. For a directed graph with a number of nodes, the corresponding joint probability is

$$p(x) = \prod_{k=1}^{K} p(x_k | pa_k). \qquad (2)$$

Among them, pa_k represents the parent node of the meaning.

When multiple nodes have similar probability distributions, such as the same parent and child nodes, we can express this in a more concise way. For example, the probability distribution is

$$p(t, w) = p(w) \prod_{k=n}^{N} p(t_n | w). \qquad (3)$$

It will be more convenient if the parameters of the model can also be expressed. Consider a probabilistic model with parameters:

$$p(t, w | x, \alpha, \sigma^2) = p(w | \alpha) \prod_{k=n}^{N} p(t_n | w, x_n, \sigma^2). \qquad (4)$$

We introduce the likelihood function in Bayesian networks. The likelihood function is a function representing the model parameters. When the samples in D are independent and identically distributed, we can get

$$L(w) = p(x_1, x_2, \cdots, x_N | w) = \prod_{n=1}^{N} p(x_n | w). \qquad (5)$$

We compute the negative logarithm of the likelihood function, denoting it as the error function $F(w)$:

$$F(w) = -\ln L(w) = \sum_{n=1}^{N} -\ln p(x_n | w). \qquad (6)$$

The maximum likelihood estimator can be found from this error function.

We will use the probabilistic graphical model to construct a cognitive diagnostic model suitable for computer teaching, simplify the parameter estimation according to the d-partition, and use the likelihood function to calculate the error function of the model to obtain a parameter optimization algorithm.

3.2. Active Learning. Active learning is a type of semisupervised machine learning that is aimed at reducing the cost of model training in machine learning. By prioritizing uncertain samples, the model allows experts (manual) to focus on providing the most useful information. This helps the model learn faster and allows experts to skip data that is not very helpful to the model. In some cases, this can greatly reduce the number of labels that need to be collected from experts and still get a good model. This can save time and money for machine learning projects. In supervised learning, thousands of labeled instances are required to train the model. Sometimes, the cost of these tags is small, even negligible, such as tags for spam and reviews for movies.

There are three main frameworks for active learning, and Figure 3 shows the differences between the three frameworks.

The first is the membership synthesis query, in which the learner can request the label of any unlabeled instance or even generate the instance himself to request the label. This approach works well for many problems, but if the labeler is a human, it may generate instances that human labelers cannot.

The second framework is stream-based selective sampling. In this method, instances are given to the model in a certain order, and the model chooses whether to label the sample.

The last is pool-based sampling. This approach assumes the existence of a small amount of labeled data and a large amount of unlabeled data and usually utilizes some strategy to selectively pick instance queries from a pool of unlabeled data in a greedy manner.

3.3. Cognitive Diagnosis Model. The DINA model (deterministic inputs, noisy, and gate model) is a typical discrete cognitive diagnosis model. The model describes students as a multidimensional knowledge point mastery vector and diagnoses them from the actual answer results of students. Because DINA model is simple, the interpretability of parameters is good, and the complexity of DINA model is not affected by the number of attributes. In the DINA model, there are M students, N questions, and K knowledge points. Use $R_j I$ to represent the jth student's answer to the i-th question, where $j = 1, 2M, i = 1, 2 \cdots N$. When $R_j = 1$, it
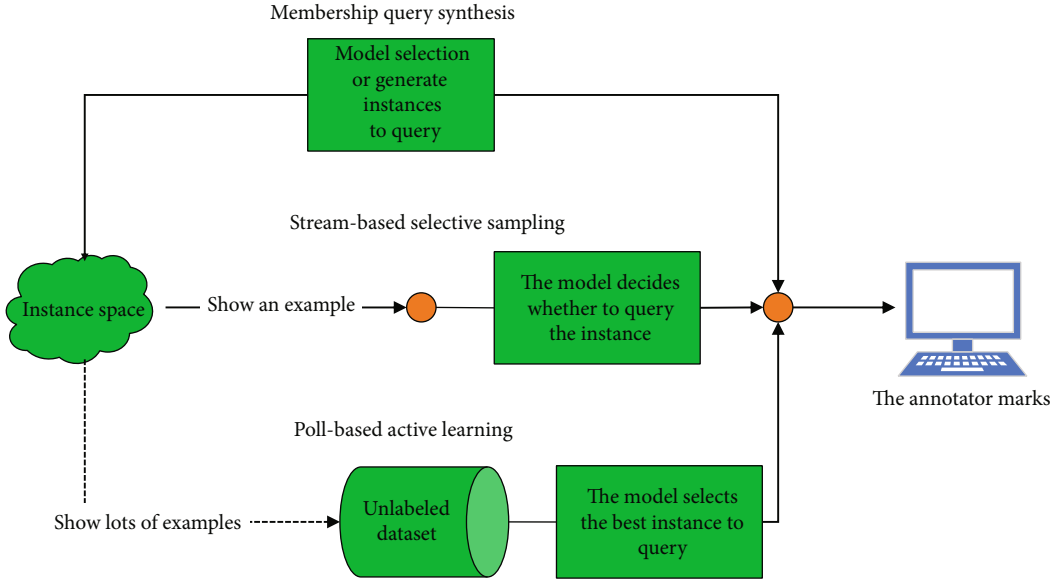
Membership query synthesis
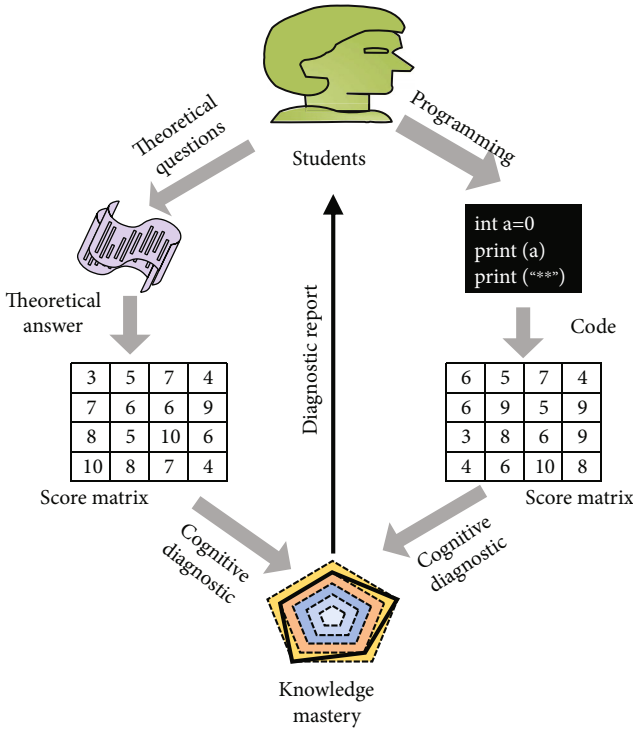

FIGURE 3: Three main active learning frameworks.

FIGURE 4: The specific application process of cognitive diagnosis in the teaching of computer courses.

means that the answer is correct, and $R_i = 0$ means that the answer is wrong. $9xi$ indicates whether the i-th question examines the k-th knowledge, $9ki = 1$ means that answering the question i requires knowledge point k, and $9xi = 0$ means no need, where $k = 1, 2 \cdots K$. $a_i k$ means the k-th whether i students have mastered the k-th knowledge point, $a_i k = 1$ means mastered, $a_i k = 0$ means no mastery. The potential

answers of student j on question i are

$$\eta_{ji} = \prod_{k=1}^{K} \alpha_{jk}^{q_{ki}}. \tag{7}$$

If $n;;= 1$, it means that the student answered the question correctly. The DINA model assumes that students need to master all the knowledge points examined by the question in order to answer the question correctly; if $n;i = 0$, that is, the student did not answer the question correctly, it means that student j has not mastered at least one of the knowledge examined by the question i.

In the DINA model, the Q matrix is used to represent the relationship between topics and knowledge points, and it is assumed that each knowledge point is independent. We will discuss in Section 5 how to incorporate preconditional relationships between knowledge points into the model.

In addition, DINA introduces two test question parameters: slip and guess. sq is slip, which means the probability that the student has mastered all the knowledge points examined by the question i but made a mistake and the probability that the student has completely mastered all the knowledge points examined by the question i but answered correctly. Therefore, in the model DINA, the probability of a student answering the question correctly is

$$p\left(R_{ji} = 1 \mid \alpha_j\right) = g_i^{1-\eta_{ji}} \left(1 - s_i\right)^{\eta_{ji}}. \tag{8}$$

DINA assumes that students' responses to each question are independent, so the conditional distribution of student j 's score is

$$L\left(R_j \mid \alpha_j\right) = \prod_{i=1}^{N} p_i\left(\alpha_j\right)^{R_{ji}} \left(1 - p_i\left(\alpha_j\right)\right)^{1-R_{ji}}. \tag{9}$$
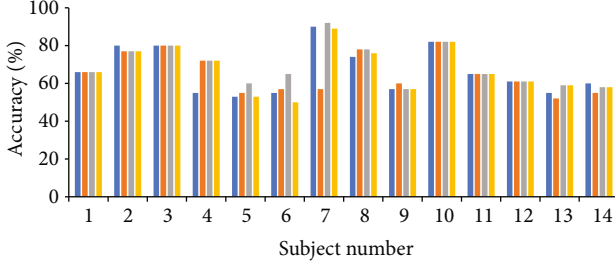
Figure 5: Dataset accuracy.

For M students, the conditional distribution of all scores A is

$$L(R|\alpha) = \prod_{j=1}^{M} L(R_j|\alpha_j) = \prod_{j=1}^{M} \prod_{i=1}^{N} p_i(\alpha_j)^{R_{ji}} (1 - p_i(\alpha_j))^{1-R_{ji}}. \tag{10}$$

Assume

$$\widehat{v} = \left(\widehat{s_1}, \widehat{g_1}, \cdots, \widehat{s_N}, \widehat{g_N}\right). \tag{11}$$

The marginal likelihood of the student's score is

$$L(R) = \prod_{j=1}^{M} L(R_j) = \prod_{j=1}^{M} \prod_{l=1}^{L} L(R_j|\alpha_l) p(\alpha_l), \tag{12}$$

where $L = 2k$. Formula (12) can also be expressed as

$$p_i(\alpha_l) = \begin{cases} g_i, & \text{if } \alpha_l^T q_i < q_i^T q_i, \\ 1 - s_i, & \text{if } \alpha_l^T q_i = q_i^T q_i. \end{cases} \tag{13}$$

Then, the log-likelihood function of the score and distribution is

$$F(R) = \log \prod_{j=1}^{M} L(R_j) = \sum_{j=1}^{M} \log(L(R_j)). \tag{14}$$

Derivative with respect to v, we get

$$\begin{aligned}
\frac{\partial F(R)}{\partial v_{i\eta}} &= \frac{\partial g_i}{\partial v_{i\eta}} \left(\frac{1}{g_i(1-g_i)}\right) \sum_{\{\alpha_i : \alpha_1^T q_i < q_i^T q_i\}} (\xi_{il} - g_i \varepsilon_l) \\
&\quad + \frac{\partial(1-s_i)}{\partial v_{i\eta}} \left(\frac{1}{s_i(1-s_i)}\right) \sum_{\{\alpha_i : \alpha_1^T q_i = q_i^T q_i\}} (\xi_{il} - (1-s_i)\varepsilon_l) \\
&= \frac{\partial g_i}{\partial v_{i_n}} \left(\frac{1}{g_i(1-g_i)}\right) \left(\xi_{il}^{(0)} - g_i \varepsilon_l^{(0)}\right) \\
&\quad + \frac{\partial(1-s_i)}{\partial v_{i\eta}} \left(\frac{1}{s_i(1-s_i)}\right) \left(\xi_{il}^{(1)} - (1-s_i)\varepsilon_l^{(1)}\right),
\end{aligned} \tag{15}$$

in

$$\varepsilon_{il} = \sum_{j=1}^{M} p(\alpha_l|R_j). \tag{16}$$

Therefore, we can get

$$p(\alpha_l|R_j) = \frac{p(R_j|\alpha_l) p(\alpha_l)}{\sum_{l=1}^{L} p(R_j|\alpha_l) p(\alpha_l)}, \tag{17}$$

$$\widehat{s_i} = \frac{\varepsilon_{il}^{(1)} - \xi_{il}^{(1)}}{\varepsilon_{il}^{(1)}}, \tag{18}$$

$$\widehat{g_i} = \frac{\xi_{il}^{(0)}}{\varepsilon_{il}^{(0)}}. \tag{19}$$

The parameters can be iteratively optimized according to the above three formulas.

4. Experimental Results and Analysis

In the existing cognitive diagnosis model, students' proficiency in using skills or knowledge points refers to their ability to use this knowledge to solve problems. In this article, we call it theoretical learning ability. In the field of computer teaching, only diagnosing students' theoretical learning ability cannot meet the teaching needs. We also need to consider students' ability to convert theoretical knowledge into code. Therefore, the traditional cognitive diagnosis model has been unable to meet the needs of computer teaching. To address this issue, we propose a new framework for computer science teaching: CDF-CSE.

The specific application process of cognitive diagnosis in computer course teaching is shown in Figure 4. First, students pass exams or assignments, answer theoretical questions, or write code; teachers or teaching assistants give a specific score based on students' answers; cognitive diagnostic models make inferences about students' abilities based on these scores, as well as some other teaching information, generate a cognitive report, and finally return it to the student. Students can carry out targeted training based on the cognitive report to check and fill in the gaps.

To validate the proposed model, we collected classroom data from a computer course at the University of Science and Technology of China to validate our model. We organized, cleaned, and formatted the collected real-world datasets, excluding some special data, such as the data of students who barely handed in homework. Our experiments were conducted on two real datasets (data from the computer courses "data structure" and "network security") and one simulated dataset, all three datasets containing students' theoretical questions and lab grades R, R', as well as the questions and the knowledge Q, Q' examined in the experiment; in the real dataset, the students' grades and the knowledge points required by the questions are provided by teachers or teaching assistants. The base case of the dataset is presented in Figure 5.

TABLE 1: Experimental results for predicting student performance.

| Model	Data structure		Network safe		Simulation dataset	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
IRT	0.358 ± 0.02	0.415 ± 0.02	0.203 ± 0.02	0.247 ± 0.02	0.227 ± 0.02	0.327 ± 0.02
PMF-5D	0.316 ± 0.02	0.380 ± 0.02	0.195 ± 0.02	0.246 ± 0.02	0.212 ± 0.02	0.265 ± 0.02
PMF-10D	0.319 ± 0.02	0.383 ± 0.02	0.196 ± 0.02	0.248 ± 0.02	0.219 ± 0.02	0.265 ± 0.02
PMF-KD	0.319 ± 0.02	0.383 ± 0.02	0.196 ± 0.02	0.248 ± 0.02	0.213 ± 0.02	0.265 ± 0.02
FuzzyCDF	0.283 ± 0.02	0.376 ± 0.02	0.210 ± 0.02	0.230 ± 0.02	0.220 ± 0.02	0.314 ± 0.02
NeuralCD	0.464 ± 0.02	0.481 ± 0.02	0.429 ± 0.02	0.498 ± 0.02	0.120 ± 0.02	0.235 ± 0.02
CDF-CSE	0.147 ± 0.02	0.231 ± 0.01	0.066 ± 0.01	0.146 ± 0.01	0.106 ± 0.01	0.132 ± 0.01
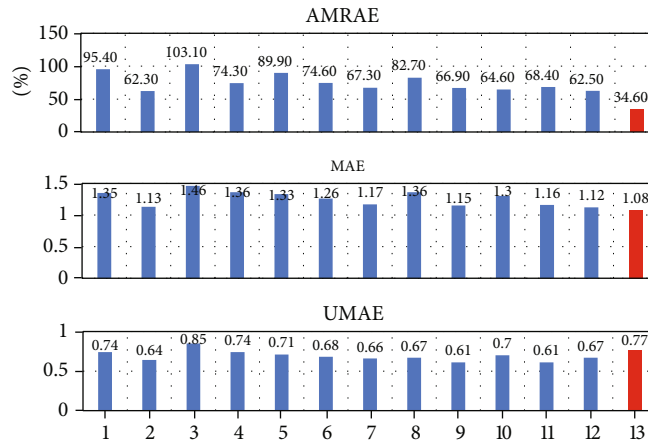
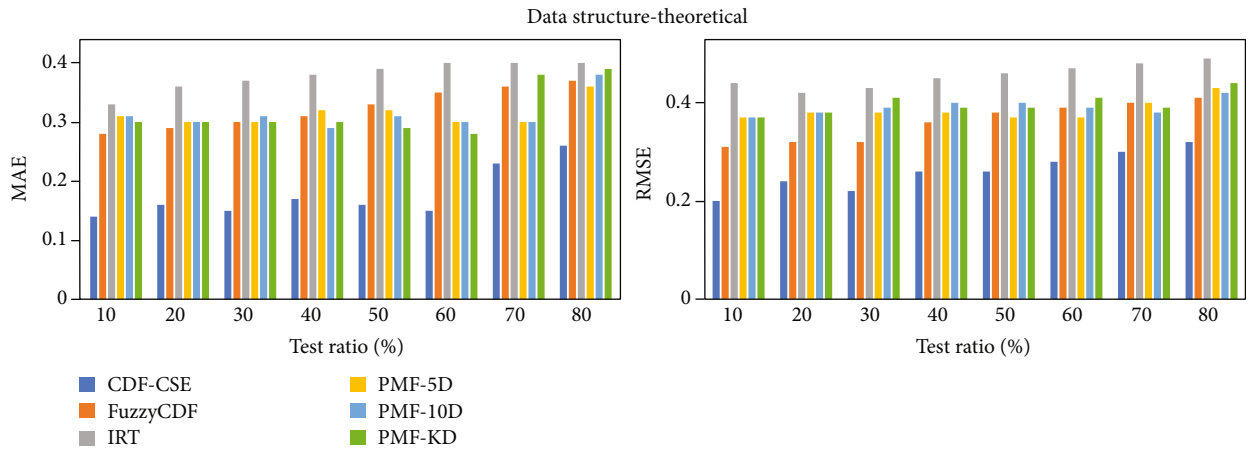FIGURE 6: Comparison of RT prediction results.

FIGURE 7: The performance of each model on the "data structure" course dataset.

First, we compare the accuracy of each model in predicting student scores to judge whether the cognitive results given by the models are reliable. The parameters a and $\#$ are the theoretical and practical abilities of the students given by the model. We use these two parameters to predict the students' scores and test the cognitive results obtained by each method through the error between the predicted scores and the actual scores. In the experiments, we use different implementations of the matrix factorization method,

namely, PMF-5D, PMF-10D, and PMD-KD, which are PMFs with 5, 10, and chi (number of knowledge points) latent factors, respectively.

First, we fixed the proportion of the training set to 80% and the test set to 20% and compared the performance of each model in the dataset containing both theoretical and experimental data. Table 1 presents the experimental results for predicting student performance. Overall, CDF-CSE performs the best on the three datasets because it can establish
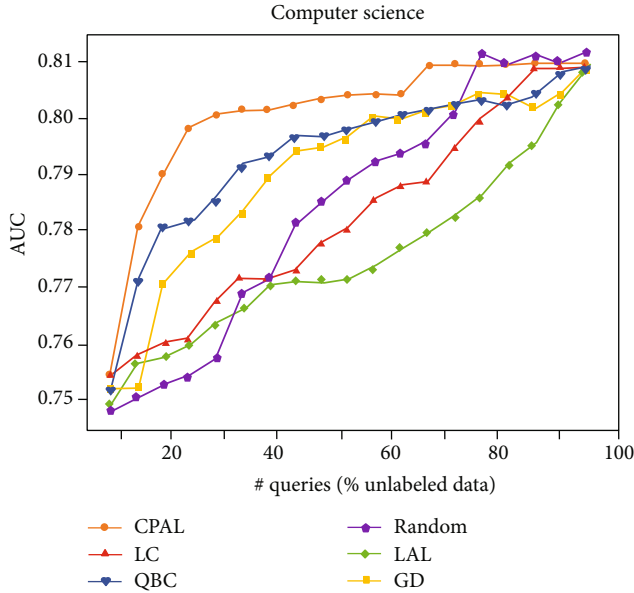
FIGURE 8: Performance of different active learning frameworks on dataset D1.
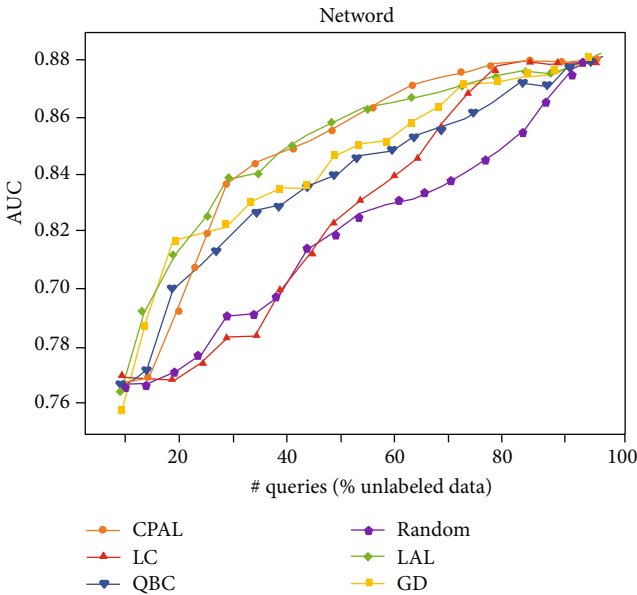

FIGURE 9: Performance of different active learning frameworks on dataset D3.

a link between theory and practice. It is worth mentioning that NeumlCD uses a deep learning method, which requires a large number of instances for training, so it performs poorly on the two smaller datasets and performs well in the larger simulated datasets, but it is still not as good as CDF-CSE.

To observe the performance of these methods on datasets with varying degrees of sparsity, we constructed training sets of different sizes (Figure 6) with 10% to 80% of each examinee's scoring data and the rest for testing. Since Neur-

alCD performs poorly in small-scale datasets, we did not compare it in this experiment. The performance of each model except NeuralCD will be compared on datasets containing only theoretical questions, datasets containing only experimental questions, and datasets containing both types. We will comprehensively compare each method from the above-mentioned different angles.

Figure 7 shows the prediction results of CDF-CSE and other methods on three different datasets. From the experimental results, we observe that CDF-CSE performs the best on all datasets.

Specifically, it outperforms PMF in terms of combining instructional assumptions, outperforming IRT in terms of quantitatively analyzing students, and outperforming all other methods in terms of combining theory and experimentation. In addition, the parameters obtained by CDF-CSE can directly represent the cognitive state of students predicted by the model, which is interpretable; however, the parameters obtained by IRT and PMF cannot give students' abilities in various knowledge points. Such a model, even if it can accurately predict students' scores, has little effect in diagnosing students' cognitive status. More importantly, as the training data becomes sparser (the proportion of training data decreases from 80% to 20%), the advantages of the CDF-CSE method gradually emerge. For example, in the comprehensive question of three datasets, when the training set is 20% and MAE is used as the judging criterion, CDF-CSE improves by 47.8%, respectively, compared with other methods that perform best on the training set, 65.8% and 49.8%.

Obviously, CDF-CSE is more accurate than other methods because the CDF-CSE method can be trained on datasets with two different problems, theoretical and experimental. That is, compared with other models that only consider one kind of problem, CDF-CSE can obtain more information during the training process; on a dataset containing two types of problems, the model will provide different probability assumptions for the two kinds of problems. This matches the actual situation. Even in the special case where the probability distribution of students' scores on both questions is the same, CDF-CSE can be trained normally; however, other models can only consider one probability distribution, which inevitably produces errors.

Figures 8 and 9 show the experimental results comparing CPAL with other strategies. The abscissa is the scale of the labeled instance, and the ordinate is the AUC result. On different computer datasets, the performance of each method is quite different. However, overall, our model outperformed the other five models on all three datasets. It can be seen that in different datasets, when the number of labeled instances is the same, our model outperforms other models most of the time. And compared with other models, our proposed active learning model can train a more accurate precondition classifier faster.

This is because CPAL selects instances according to the characteristics of knowledge point pairs, especially instances with prerequisite relationships, because the labels between them can be inferred from each other, so they are easy to be selected, while in other active learning methods, not

considering the impact of labels, they treat all instances equally in this regard. Therefore, CPAL can determine a more complete model in the early stage. The above experimental results show that CPAL can efficiently select useful instances of precondition relations. In addition, for the other five models except for CPAL, the performance of QBC is generally better, even close to CPAL on some datasets, while the performance of LAL is unstable, with good performance on some datasets and poor performance on other datasets. The worst performers are active learning methods that randomly select instances, which are not helpful in reducing training instances. Finally, we can see that the performance of most active learning models stabilizes after a certain number of labeled instances. Because there are a large number of similar unrelated knowledge points in the dataset, the amount of information they can bring is limited. Compared with the other five models, CPAL reaches the stable stage earlier, because CPAL selects the most informative instances at an earlier time.

To sum up, the use of CPAL can effectively reduce the labels required to extract computer prerequisite relations.

5. Conclusion

Microlecture is a new type of classroom teaching mode derived from the rapid development of information technology. Fragmentation and short time are the two major characteristics of microlecture teaching. It is completely different from the traditional teaching mode, but the teaching effect of the harvest is very obvious. At present, microlectures have been widely used in the teaching of schools, but in terms of practical application effects, they are not ideal. This paper mainly takes the teaching of computer courses as an example. While introducing the current situation of teaching activities, it studies the specific strategies of microlectures in computer courses, so as to better utilize the advantages of microlectures. Judging from the actual situation of the current microlecture teaching, there are still some areas to be improved and perfected, which requires teachers to change the traditional teaching concept, update the teaching method, and effectively combine the subject characteristics with the microlecture teaching. It can not only stimulate the enthusiasm and initiative of students to participate in the classroom and improve the learning efficiency but also fundamentally improve the quality of teaching, meet the educational needs of colleges, and promote the school of rapid development.

Data Availability

The figures and tables used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] H. Chaussade, X. Cazals, G. Desoubeaux et al., "Central nervous system candidiasis beyond neonates: lessons from a nationwide study," *Medical Mycology*, vol. 59, no. 3, 2021.

[2] M. Cents-Boonstra, A. Lichtwarck-Aschoff, M. M. Lara, and E. Denessen, "Patterns of motivating teaching behaviour and student engagement: a microanalytic approach," *European Journal of Psychology of Education*, vol. 37, no. 1, pp. 227–255, 2021.

[3] M. R. Msimanga, "The impact of micro teaching lessons on teacher professional skills: some reflections from South African student teachers," *International Journal of Higher Education*, vol. 10, no. 2, p. 164, 2020.

[4] M. Nasir, M. Rockmore, and C. M. Tan, "Do the lessons from micro-conflict literature transfer to high crime areas?: examining Mexico's war on drugs," *Journal of Development Studies*, vol. 56, no. 1, pp. 26–44, 2020.

[5] E. Kolbachev, E. Sidorova, and P. Vaneeva, "Human factor and cognitive methods in the design of products and production systems of mechanical engineering in the framework of NBIC convergence," *International Conference on Intelligent Human Systems Integration*, vol. 1131, no. 3, pp. 258–272, 2020.

[6] M. Roldán-Palacios and A. López-López, "Disfluency as an indicator of cognitive-communication disorder through learning methods," in *International Conference on Brain Informatics*, Springer, Cham, 2021.

[7] N. J. Hodges, P. A. Wyder-Hodge, S. Hetherington, J. Baker, Z. Besler, and M. Spering, "Topical review: perceptual-cognitive skills, methods, and skill-based comparisons in interceptive sports," *Optometry and Vision Science*, vol. 98, no. 7, pp. 681–695, 2021.

[8] K. A. Christensen, *Elementary Students as Transpersonal Researchers: Guided Cognitive Imagery on Science Test Scores Using a Mixed-Methods Approach*, 2020, Ph.D. Thesis.

[9] C. Kaufman, "Civic education in a fake news era: lessons for the methods classroom," *Journal of Political Science Education*, vol. 17, no. 2, 2021.

[10] F. Betsy and M. Shruti, "Impact of technological developments on infectious disease epidemiology: lessons from the first 100 years of AJE," *American Journal of Epidemiology*, vol. 13, no. 2, 2022.

[11] C. F. Collares, "Cognitive diagnostic modelling in healthcare professions education: an eye-opener," *Advances in Health Sciences Education*, vol. 27, no. 2, pp. 427–440, 2022.

[12] P. Horopaha and Pavliuk, "Formation of pupil's ecological competence with speech disorders at mathematics lessons," *Scientific bulletin of KRHPA*, vol. 13, no. 13, pp. 128–140, 2020.

[13] A. Zendler and H. Greiner, "The effect of two instructional methods on learning outcome in chemistry education: the experiment method and computer simulation," *Education for Chemical Engineers*, vol. 30, no. 13, pp. 9–19, 2020.

[14] E. Atasoy and A. Bak, "Investigation of students' cognitive learning in mathematics lessons supported with writing activities," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 20, no. 3, pp. 335–345, 2020.

[15] J. Bosica, "Using a mixed methods approach to study the relationship between mathematics anxiety, mathematics teacher efficacy, and mathematics teaching anxiety in preservice elementary school teachers in Ontario," *Canadian Journal of Science, Mathematics and Technology Education*, vol. 22, no. 1, pp. 190–209, 2022.

[16] Y. J. I. Chen, Y. H. Chen, J. L. Anthony, and N. A. Erazo, "Evaluation of the computer-based orthographic processing assessment: an application of cognitive diagnostic modeling," *Journal of Psychoeducational Assessment*, vol. 40, no. 2, pp. 271–292, 2022.

[17] D. A. Nikolenko, V. D. Munister, A. L. Zolkin, I. L. Khvatov, and A. A. Karapetyan, "Analysis of modern computer diagnostic and measurement methods in medicine," *Journal of Physics Conference Series*, vol. 1889, no. 5, article 052005, 2021.

[18] N. E. Barrett, W. C. Hsu, G. Z. Liu, H. C. Wang, and C. Yin, "Computer-supported collaboration and written communication: tools, methods, and approaches for second language learners in higher education," *Human Behavior and Emerging Technologies*, vol. 3, no. 2, pp. 261–270, 2021.

[19] K. Syed, W. C. Sleeman, J. J. Nalluri et al., "Artificial intelligence methods in computer-aided diagnostic tools and decision support analytics for clinical informatics," *Artificial Intelligence in Precision Health*, vol. 19, no. 3, pp. 31–59, 2020.

[20] W. Wang, L. Song, T. Wang, P. Gao, and J. Xiong, "A note on the relationship of the Shannon entropy procedure and the Jensen–Shannon divergence in cognitive diagnostic computerized adaptive testing," *SAGE Open*, vol. 10, no. 1, article 215824401989904, 2020.