International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 667-677

https://doi.org/10.70517/ijhsa46460

Research on the Construction of Online Learning Evaluation System Based on Bayesian Networks and Multi-Dimensional Evaluation Methods

Yunyi Zhu[1,*] and Xinting Yue[1]
[1] Zhejiang Tongji Vocational College of Science and Technology, Hangzhou, Zhejiang, 310000, China
Corresponding authors: (e-mail: z20220150801@zjtongji.edu.cn).

Abstract Traditional learning assessment methods rely on static tests, which cannot reflect students' learning status in real time. In this paper, we propose an online learning assessment system based on Bayesian network, which can dynamically assess the learning effect of students and update their knowledge status in real time. First, the system collects data from the online learning platform through Python crawler technology, including learners' test scores, homework scores and learning behavior data. Then, a Bayesian network model is used to model the learning process of the students, assess their knowledge mastery, and calculate the probability of answering questions correctly by combining with IRT theory. Through experimental validation, this system performs well in terms of assessment accuracy and prediction accuracy, with an assessment accuracy of 92.65% and a prediction accuracy of 90.84%. In addition, the system is able to track learners' behavioral characteristics in real time and improve the effect of personalized teaching by analyzing learners' learning patterns. The experimental results show that the online learning assessment model based on Bayesian network can effectively improve the accuracy of learning assessment and provide an efficient learning assessment tool for online education platforms.

Index Terms Bayesian network, online learning, learning assessment, data collection, learning behavior, accuracy rate

I. Introduction

With the popularization of the Internet and the progress of technology, online education has become a flexible and convenient way of learning, which has been widely used and promoted, especially after the epidemic was widely used [1]. The comprehensive development of online education makes it possible for people to learn directly at any location, reducing the cost of space. And it supports post-course commitment and repeated focus learning, which makes access to resources more convenient and promotes the development of fair education. You can listen to lectures directly through web pages, apps and other related tools, so that students can freely choose the learning content that best suits them and innovate their learning style [2]-[5]. However, the traditional online education model still has some problems, such as insufficient teacher resources, low student participation, and inefficient assessment [6]. Among them, online learning assessment is not only an effective means to measure students' learning effectiveness, but also to improve the retention rate of platform users [7].

However, at present, the nature of online learning assessment is more or less the same, which is based on the assessment method of explicit feedback of assignments and online hours. This approach is commonly used in offline teaching. However, when this method is copied to online learning, there are many problems, such as students can independently decide the time to complete the homework, can choose the way to complete the homework or even through Baidu and other tools to assist themselves in completing the homework, which results in the completion of the homework can not be a very objective reflection of the student's learning effect [8]-[10].

In addition, other online learning assessment strategies have a serious lag, lack of personalization, ignore students' behavioral data in multiple dimensions, and are not conducive to dynamic real-time assessment, which makes it difficult for students and teachers to identify learning problems in time, leading to unsatisfactory online learning results [11]-[14]. In this context, Bayesian network brings new opportunities and challenges for online education. Bayesian network is a probabilistic graphical model that describes the dependencies between variables and can be used to infer unobserved data [15].

With the rapid development of information technology, online learning has become an important trend in education. However, traditional learning assessment methods often fail to track students' learning progress in real time and lack flexibility and personalization in the assessment process. In order to adapt to this trend, researchers have gradually proposed intelligent assessment systems based on technological means. Bayesian network, as an excellent reasoning tool, can realize dynamic reasoning for complex systems by modeling the relationship between various variables. In online learning, the application of Bayesian network can provide real-time and accurate learning assessment results to help teachers understand students' knowledge mastery in a timely manner and optimize teaching programs.

Bayesian networks utilize graphs to represent the dependency relationship between knowledge points, and calculate students' mastery of a certain knowledge point through conditional probability tables. Unlike the traditional assessment method based on a fixed question bank, Bayesian network can speculate the knowledge level of students in real time based on their specific answers and historical data, adapting to the personalized learning needs of students. In online education platforms, students' learning behavior data can provide an important basis for assessment, and using these data, Bayesian networks can build accurate learning models and achieve more dynamic learning assessment.

In addition, with the advancement of data collection technology, learning platforms are able to obtain students' learning behavior data in real time, such as the length of video viewing and the number of test answers. These behavioral data are important for assessing students' learning process and results. By integrating learning behavior data and Bayesian network inference models, real-time learning feedback can be provided to students, thus providing educators with personalized teaching strategies. The research in this paper not only enhances the potential of Bayesian networks in education, but also provides a new idea and method for online learning assessment. This paper includes three main aspects: data collection, model construction and effect evaluation. In terms of data collection, Python crawler technology is used to collect students' learning behavior data and their test scores from online learning platforms, and the data are cleaned and standardized. In terms of model construction, Bayesian networks were used to model students' knowledge mastery, and IRT theory was combined to determine the correct probability of students' answers to questions. The inference ability of the model is realized through real-time analysis of student behavioral data. In terms of effect evaluation, the accuracy of the model is verified through experiments, and its application effect in a real online learning environment is evaluated, mainly examining the assessment accuracy, prediction accuracy, and error, which ultimately provides decision support for personalized learning.

II. Bayesian networks

Bayesian networks [16], [17] can graphically express the probabilistic relationships between network variables and consist of a directed acyclic graph (DAG) and a conditional probability table (CPT). The DAG is the structure of a Bayesian network that graphically visualizes and qualitatively shows the dependencies between variables and consists of nodes and directed edges. The nodes represent the variables and the directed edges connecting the nodes represent the dependencies between the variables. The starting point of the directed edge is the parent node, and the end point is the child node. Generally speaking, there is a causal relationship between the parent node and the child node, i.e., the parent node is the “cause” and the child node is the “effect”. If a node does not have a parent, it is an evidence node. If a node has no children, it is the target node.CPT quantitatively describes the dependency relationship between variables by incorporating probability theory. The related computational principles involved are as follows:

(1) A priori probability: refers to the probability of each variable occurring before there is no new observational evidence, as determined by methods such as past subjective experience or historical data analysis, and is set as $P(X_i)$.

(2) Conditional probability: refers to the probability of the occurrence of another variable under the condition that a variable is known to have occurred, reflecting the dependency relationship between the variables of the Bayesian network, which can be determined by historical data or subjective experience. If node X_i points to node $pa(X_i)$ through a directed edge, then node X_i is said to be the parent of $pa(X_i)$ and $pa(X_i)$ is the child of X_i. When $P(X_i) > 0$, then the probability that its parent node $pa(X_i)$ occurs in the case of node X_i is:

$$P(X_i \mid pa(X_i)) = \frac{P(X_i \cap pa(X_i))}{P(pa(X_i))} \tag{1}$$

(3) Joint probability: refers to the probability of simultaneous occurrence of multiple variables in a Bayesian network, which is calculated based on the node dependencies and conditional probabilities in the network. In a Bayesian network, the joint probability can be obtained by multiplying the product of the conditional probabilities of each node, and the joint probability distribution of all the nodes in the Bayesian network is calculated as follows with the known prior probability of the node and the conditional probability of the other child nodes:

$$P(X)=P(X_1,X_2,\cdots,X_n)=\prod_{i=1}^{n}P(X_i\mid pa(X_i)) \tag{2}$$

(4) Bayes formula: if the probability of occurrence of both event X_i and its parent node $pa(X_i)$ is greater than 0, then $P(X_i \mid pa(X_i) \mid pa(X_i))P(pa(X_i)) = P(pa(X_i) \mid X_i)P(X_i)$, and derive after which the Bayesian formula can be obtained:

$$P(X_i\mid pa(X_i))=\frac{P(pa(X_i)\mid X_i)P(X_i)}{P(pa(X_i))} \tag{3}$$

(5) Posterior probability: refers to the update of the probability of occurrence of a consequence variable after the network obtains new evidence, this update is based on Bayes' theorem, which is formulated as follows:

$$P(X_i\mid pa(X_i))=\frac{P(pa(X_i)\mid X_j)P(X_i)}{\sum_{k=1}^{n}P(pa(X_k)\mid X_k)P(X_k)} \tag{4}$$

where $p(X_i \mid pa(X_i))$ is the posterior probability of the event X_i .

III. Bayesian network-based learning assessment models

This section focuses on constructing an approach to Bayesian network online learning assessment modeling. The student model is described by the overlay student model, which considers the student's knowledge as a subset of the expert's knowledge, and is implemented with Bayesian network techniques, which results in an assessment model with strong inference and prediction capabilities.

III. A. Relevant Variables

III. A. 1) Knowledge variables

In order to assess students' knowledge, different variables are used to represent knowledge at different levels of granularity. These variables can represent knowledge in multiple senses. For simplicity, these different variables are called concepts, topics and subjects.

Concepts: are basic units of knowledge called meta-knowledge items that cannot be further decomposed.

A T_i is a (C^i, W) pair:

where C^i is the set of mutually independent basic concepts, $C^i = \{C_{i1}, C_{i2}, \cdots, C_{in}\}$.

$w = (w_{i1}, \cdots, w_{in})$ is a vector of specified weights indicating the relative importance of each concept with respect to the TOPIC to which it belongs.

A random variable T_i is used to denote a student's knowledge about a topic, T_i is defined as:

$$T_i=\sum_{j=1}^{n_i}W_{ij}C_{ij} \tag{5}$$

A discipline is a (T, α) pair:

T is a collection of mutually independent topics $T = \{T_1, T_2, \cdots, T_n\}$.

$\alpha = \{\alpha_1, \cdots, \alpha_n\}$ is a vector of specific weights indicating the relative importance of each topic for the discipline it belongs to. To represent the knowledge of a subject A , a random variable A is represented as in Eq:

$$A=\sum_{i=1}^{s}\alpha_i T_i \tag{6}$$

III. A. 2) Evidentiary variables

These variables are used to collect information about the learner, which can be questions, tasks, multiple choice questions, etc., as long as the system has the ability to evaluate whether the learner answered correctly or not. A random variable Q obeying a Bernoulli distribution is used to represent a question, where $Q=1$ means that the learner answered question Q correctly and vice versa $Q=0$. The probability distribution of Q is given in Eq:

$$P(Q=x)=p^x(1-p)^{(1-x)} \tag{7}$$

where $p = P(Q=1)$, $x \in \{0,1\}$.

III. B. Connection relations and parameters

III. B. 1) Aggregate relationships between knowledge items

In order to discuss the relationship between variables, suppose a knowledge item I is divided into finite terms $I_1, I_2, \cdots, I_n$. In order to model the relationship between them. There are two possible models: 'Model 1' denotes that mastery of $I_1, I_2, \cdots, I_n$ has a causal effect on mastery of I, and 'Model 2' denotes that mastery of I has a causal effect on mastery of $I_1, I_2, \cdots, I_n$ has a causal effect.

In this paper, we choose model 1. model 1 well models the incremental learning process of a learner, i.e., in order to learn a topic, one must learn all the concepts that make up this topic. So the structure of BN in this paper is shown in Fig. 1.

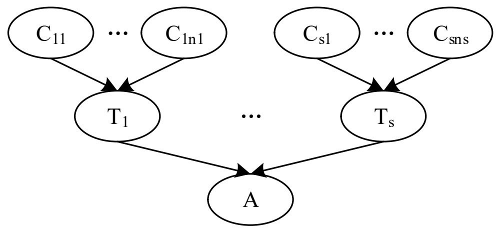

Figure 1: BN structure drawing

III. B. 2) Parameters between knowledge items

In order to reduce the parameters of the Bayesian network and thus reduce the time for updating the network, in this paper, only two states are considered for the composite knowledge items, Master and UMasetr, which represent mastery and non-mastery, respectively.

In this paper, the students' mastery level of the meta-knowledge item is categorized into four levels i.e., High, Medium, Low, and UMaster, the meaning of which is shown in Eq:

$$\theta_{I_i} = \begin{cases} High & 0.8333 < P(I_i = Master) < 1 \\ Medium & 0.5 < P(I_i = Master) \leq 0.8333 \\ Low & 0.167 < P(I_i = Master) \leq 0.5 \\ UMaster & 0 < P(I_i = Master) \leq 0.167 \end{cases} \tag{8}$$

where θ_{I_i} represents the students' mastery level of the knowledge item I_i, High means that the students have mastered the knowledge item and are able to apply it in practice, Medium means that the students' mastery level of the knowledge item is medium, basically mastering the knowledge item, but not able to apply the knowledge proficiently, and need to learn more, Low means that the students' mastery level of the knowledge item is relatively poor, and still need to strengthen the learning, UMaster means that the students have not mastered the knowledge item, and need to review the preparatory knowledge of the knowledge item. Low means that the students' mastery of the knowledge item is poor, and still need to strengthen the learning; UMaster means that the students have not mastered the knowledge item, and need to review the preparatory knowledge of the knowledge item. The specific conditional probability rules between nodes in this paper are given:

(1) If the sub-knowledge item I_i is a meta-knowledge item, i.e., type C, and it is known that the weight of I_i on its parent knowledge item is w_i, then the weight of I_i in various cases of taking the value $\theta_{I_i} \in \{High, Medium, Low, UMaster\}$ is as follows:

$$w_{i'} = \begin{cases} w_i & if \;\; \theta_{I_i} = High \\ \frac{2}{3} w_i & if \;\; \theta_{I_i} = Medium \\ \frac{1}{3} w_i & if \;\; \theta_{I_i} = Low \\ 0 & if \;\; \theta_{I_i} = UMaster \end{cases} \tag{9}$$

(2) If the child knowledge item I_i is a composite knowledge item i.e., T or A type, and the weight of I_i on its parent knowledge item is known to be w_i, then the weight of this knowledge item $w_i' = w_i$ for various values of values (Master, UMaster).

(3) The conditional probability distribution of the parent knowledge item I is shown in Eq:

$$P(I / I_1, \cdots, I_s) = \sum_{i \in s} w_i' \tag{10}$$

where s is the number of subknowledge items.

III. B. 3) Parameters between evidence nodes and knowledge nodes

The parameters that need to be specified in the evidence nodes and knowledge nodes are the prior probability of CONCEPTS and the conditional probability of the evidence nodes given the CONCEPTS condition. In this paper, a three-parameter model based on IRT theory is used to determine the conditional probability, which is mathematically modeled as Eq:

$$G(x) = c + \frac{(1-c)}{1+\exp\left(-1.7a(x-b)\right)} \quad x \ge 0 \tag{11}$$

where a denotes the differentiation, b denotes the difficulty level of the test questions, and $c = 1/n$ denotes the guessing coefficient, the difficulty coefficient and differentiation of the test questions are analyzed and given by the experts or teachers.

The function $G(x)$ is used to calculate the probability of answering a question correctly under the condition that the student is known to have mastered the concepts. In this paper it is calculated as follows: if a student has not mastered any of the relevant concepts, the probability of answering correctly is set to $c = 1/n$. If all the relevant concepts are mastered, the probability of answering correctly is set to $1-s$, with s being the miss factor, which indicates the probability of answering the test question incorrectly due to some chance factor. The rest of the probabilities were calculated using the function $G(x)$, between $1/n$ and $1-s$.

The method of calculating the remaining probabilities using the function $G(x)$ is described as follows: let x^* denote $G(x^*) = 1-s$ and assume that there are p concepts associated with a particular test item. Then the values being used to compute 2^p probabilities are as follows:

$$\left\{ G(0), G\left(\frac{x^*}{2^p-1}\right), G\left(\frac{2x^*}{2^p-1}\right), \cdots, G\left(\frac{(2^p-2)x^*}{2^p-1}\right), G(x^*) \right\} \tag{12}$$

III. C. Bayesian Networks Evaluation Model for Online Learning

Synthesizing the above theories to get the Bayesian network online learning assessment model schematic shown in Figure 2. In the diagram, A node can represent a course, T can represent a chapter, C can represent a knowledge point in a certain chapter, and E represents the test questions, which are used to verify whether the learner has mastered the relevant knowledge and the level of mastery hierarchy. When the system builds the test bank, the test questions are organized in the order of knowledge points, and each question is cognitively transformed. In addition to the test question ID, question stem, and options, each question is labeled with the sequence number of the knowledge point examined, the cognitive ability tested by the question, the coefficient of difficulty, the degree of differentiation, and the correct answer. The assessment process was divided into two phases:

a. Diagnostic phase: the diagnostic phase is accomplished by the bottom layer of the network, i.e., the part of the network consisting of the CONCEPTS and QUESTIONS nodes. This phase is accomplished by the testing process, which assesses what the students have mastered through their answers. The process of diagnosis is the decision reasoning process, which is accomplished in this paper by utilizing the powerful reasoning mechanism of Bayesian networks.

b. Evaluation stage: Based on the results of the evaluation stage, the probability propagation of Bayesian network achieves the evaluation of students' knowledge level at different granularity levels and real-time updating of students' latest knowledge status level.

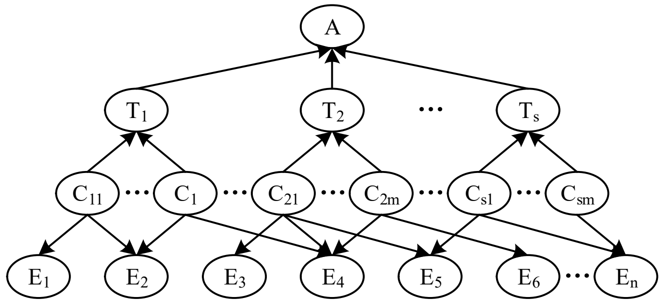

Figure 2: The BN online learning evaluation model schematic

IV. Multidimensional assessment based on an online learning assessment system

IV. A. Online Learning Data Acquisition and Processing

In this chapter, a new learning data set (MADS) is constructed by using crawler technology to collect and process the relevant learning data of courses in “an online learning platform”.

IV. A. 1) Data acquisition

In this study, a number of courses of “an online learning platform” are selected as the target of data collection. The platform is rich in teaching resources, has a large number of elective learners, and the degree of completion of course tasks by learners is relatively high, so the data quality is relatively good. However, there are many dimensions of course-related data, and the test questions answered by each learner in the exam are randomly selected from the question bank, which bring great pressure to the data statistics work, so we choose to use Python crawler technology to collect learners' chapter quiz scores, homework scores, and midterm and final exam answer records.

When collecting data based on Python crawler technology, firstly, use selenium to open the official website of “an online learning platform”, log in to the account and enter the background management interface. Enter into the teaching content release - midterm/final exams, use beautiful soup to build a document tree, and save the questions and the corresponding stem information in the form of a dictionary. Then enter the student performance management and cycle through the operations to collect the data of students' online learning behavior, homework scores, quiz scores, exam answer information and answer results, and so on.

IV. A. 2) Data pre-processing

Learners' learning data of several courses of “an online learning platform” are collected by crawler technology. After processing, a learning dataset containing 1375 learners is constructed. The dataset includes learner ids, online learning behaviors, effort characteristic parameters, test score records, and test question-knowledge point examination matrix. Among them, the learners' effort feature parameters are the total scores of the assignments and quizzes for the learners. Because the assignments and quizzes are of low difficulty, the learners' grades for these parts basically reflect the level of effort they have put in the course learning process. The questions that learners answer are randomly selected from a bank of 100 questions. The questions in the question bank were of graded difficulty, comprehensive in their examination of knowledge, and distinguished significantly between learners' levels of knowledge. The test question score record contains the learner's score for each test exercise (0 means no answer, 1 means correct answer, -1 means incorrect answer), which corresponds to the test question answer matrix. Pre-processing such as cleaning and standardization was performed on each data to reduce data noise interference.

IV. B. Online Learning Assessment System Design

The online learning effect assessment system based on Bayesian network developed in this paper mainly provides a solution to realize learning effect assessment for online learning platform. The overall design structure of the online learning assessment system is shown in Figure 3.

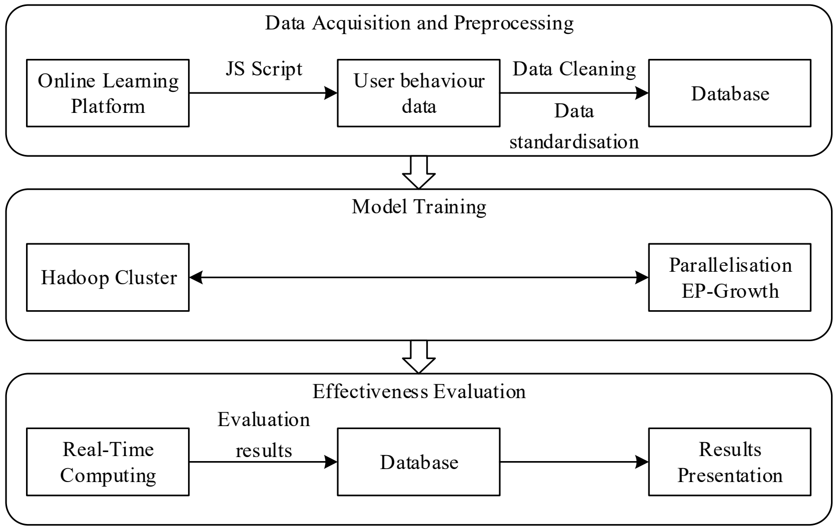

Figure 3: The overall design structure of the online learning assessment system

This system contains three parts, which are:

(1) Data Acquisition and Preprocessing Module

The main task of the data acquisition and preprocessing module is to collect real-time data about users' learning behaviors during the learning process, including the data of watching videos, operating the player, and collaborating and interacting with teachers or classmates. The collection of these data is mainly realized through the way of JS script embedding, so that the user's behavioral data can be collected without the user's perception, which improves the real-time and reliability of the data. After collecting the user's behavioral data, the background will carry out the corresponding preprocessing operations on the collected data, and the processed data will be stored in the form of logs in the database.

(2) Model Training Module

The model training module mainly mines frequent itemsets from the learning behavior data stored in HDFS, and implements Bayesian network algorithm on Spark to mine a large amount of behavior data, and eventually mines frequent itemsets of learning behaviors that represent good learning effects and frequent itemsets of learning behaviors that represent poor learning effects.

(3) Effectiveness Evaluation Module

The effect evaluation module mainly uses Sparkstreaming to evaluate the real-time learning effect of the user who is learning. The user's behavioral data in the process of learning will be stored in the form of logs in the background after collection and processing, and then Sparkstreaming will monitor the corresponding ports in real time, and the user's data will be passed into Sparkstreaming in the form of data stream. Sparkstreaming, the system will be able to calculate the user's learning effect score based on their behavioral data in real time and record it into the database. At the end of the user's learning, the platform can evaluate the user's learning effect according to his/her rating in the database, and provide specific reports and results back to the user.

IV. C. Experimental analysis of online learning assessment system

IV. C. 1) Evaluation of system performance test results

The system experiments were conducted in the open environment of the laboratory, and the testing environment was quiet and comfortable, suitable for conducting the experiments and excluding external interference from generating testing errors. Five healthy school students were selected as subjects for this system testing experiment, and the age range of the subjects was between 18 and 22 years old, including three males and two females. The subjects were required to learn the same knowledge points using the learning evaluation system designed in this paper to test the system performance. The system performance evaluation indexes include classification accuracy (Acc), total completion time, and total number of actual operations.5 subjects were required to learn the same knowledge points using the learning assessment system built in this paper to verify the performance of the system, and the results of the system performance test are shown in Table 1.

As can be seen from the data in the table, the ideal number of operations of the system performance test is 46, while the actual average number of operations is 47.2, which is caused by the misoperation of some subjects during the experiment. If misoperation occurs during the experiment, there are two ways to deal with it depending on the interface: (1) No need to correct and continue to select the correct key. (2) Add an operation to correct the error first, and then continue to select the correct key. Because of the different processing methods after the occurrence of misoperation, the number of correct operations and the total number of operations may be different for different subjects, and the average number of correct operations is 46.4. The average classification accuracy is 98.34%, and most of the misoperations of the system test occur in the interfaces of multiple buttons. In the interface of the keys, the random blinking paradigm is used and the four keys adjacent to each key with two blinks close to each other are controlled not to intersect, through this mechanism the system can still maintain a high classification accuracy rate when the number of keys is high. The average total time for subjects to complete the performance test experiment is 7.48 minutes, and the total time mainly includes command execution time, response time, command waiting time and other time. There is a large gap between the total time of different subjects, with P3 having the shortest total time of 6.83 minutes and P4 having the longest total time of 8.17 minutes. The experimental results show that the learning assessment system constructed in this paper possesses a better testing effect.

Table 1: Systematic test results

Staff	Total time (min)	Number of correct operands	Number of total operands	Acc (%)
P1	7.21	47	48	97.92
P2	7.75	46	46	100.00
P3	6.83	46	46	100.00
P4	8.17	46	47	97.87
P5	7.44	47	49	95.92
Average	7.48	46.4	47.2	98.34

IV. C. 2) Feasibility of online learning data processing

In order to verify the correctness of the Bayesian network model selected in this paper for the assessment and prediction of student performance data, the algorithm of this paper and other algorithms are deployed in “an online learning platform”, and 110758 pieces of 10 kinds of core teaching data in the platform are selected as the experimental data test set. At the same time, data evaluation accuracy, data prediction accuracy, absolute error, average error and algorithm precision were compared. The comparison algorithms are selected as Random Forest (RF), Decision Tree (DT), LSTM, and Support Vector Machine (SVM) algorithms.

The main purpose of this design is to improve the data assessment accuracy and prediction accuracy, so in order to verify the effectiveness of the proposed algorithms, these two types of accuracy are used as evaluation indexes for side-by-side comparison. The comparisons of different algorithms in terms of achievement data evaluation accuracy, prediction accuracy, absolute error, average error, and algorithm accuracy are shown in Table 2.

According to the data in the table, it can be seen that this paper's algorithm performs optimally in both data evaluation and prediction accuracy, with the two accuracy rates reaching 92.65% and 90.84%, respectively. This result is significantly better than other algorithms, the accuracy of the comparison algorithm is only 81.21% and 75.14 at the highest, and the lowest falls to 67.58% and 70.91%, which is still a more obvious gap compared with this algorithm. This shows that the high accuracy rate of the algorithm used not only demonstrates its superiority in data processing, but also provides a more reliable guarantee for practical applications.

In the process of algorithm evaluation, the absolute error and average error directly reflect the degree of deviation between the predicted and actual values, while the precision measures the accuracy of the algorithm in processing data. As can be seen from the table, this paper's algorithm has the best performance in both types of errors, which are 2.85% and 3.68%, respectively, while other algorithms have higher errors than this algorithm. In addition, this paper's algorithm significantly improves the absolute error while effectively reducing the average error, thus improving the overall performance of the algorithm. Meanwhile, in the comparison of algorithm accuracy, the accuracy of this algorithm is as high as 97.69%, which shows its excellent prediction ability. In comparison, the highest accuracy of other algorithms is only 87.84%, while the lowest is only 75.98%. This result further emphasizes the advantages of this paper's algorithm, demonstrating its reliability and accuracy in processing data.

Table 2: Comparison of data evaluation accuracy and other indexes

Algorithm	Evaluation accuracy (%)	Prediction accuracy (%)	Mean error (%)	Absolute error (%)	Precision (%)
Ours	92.65	90.84	2.85	3.68	97.69
RF	71.54	74.77	5.24	3.82	75.98
DT	67.58	70.91	6.79	6.33	77.31
LSTM	81.21	75.14	3.83	3.53	87.84
SVM	77.17	73.52	4.69	5.44	83.30

IV. C. 3) Online Learning Behavioral Characteristics Correlation

In this paper, we study the correlation between learning behaviors and total achievement through Pearson's correlation coefficient expression, mining the behavioral characteristics that are significantly correlated with learning achievement and deleting the behavioral characteristics that are weakly correlated with learning achievement. The formula is as follows:

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{13}$$

where x represents learning behavior and y represents learning performance. The online learning behavioral features of students mined by the system include: video daytime preference, video nighttime preference, test daytime preference, test nighttime preference, number of tests, video viewing duration, number of videos viewed, completing the test during break time, watching the video during the break time, test scores, and chapter scores, which are denoted in order as V1~V11. The correlation between these 11 behavioral features and the students' overall grades (S) is analyzed.

Table 3 shows the correlation coefficients between the 11 behavioral traits and academic performance. As can be seen from the table, there are six behavioral traits with eigenvalues greater than 0.4 and with correlation with academic performance, i.e., video daytime preference (V1), number of times of testing (V5), length of video viewing (V6), number of times of video viewing (V7), test scores (V10), and chapter scores (V11). These behavioral characteristics have strong correlation with learning performance, which is important for mining learners' online learning characteristics and improving learning outcomes.

Table 3: Correlation coefficient between behavior characteristics and achievement

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	S
V1	1.00											
V2	-0.40	1.00										
V3	0.03	0.05	1.00									
V4	0.07	0.04	-0.90	1.00								
V5	0.16	0.12	0.63	0.04	1.00							
V6	0.64	0.09	0.12	0.08	0.30	1.00						
V7	0.96	0.19	0.06	0.11	0.23	0.72	1.00					
V8	-0.13	-0.14	-0.03	0.05	0.01	-0.17	-0.21	1.00				
V9	0.18	0.47	0.00	0.13	0.17	0.33	0.45	0.00	1.00			
V10	0.14	0.10	0.26	0.13	0.65	0.33	0.22	-0.07	0.23	1.00		
V11	0.32	0.04	0.27	-0.11	0.29	0.39	0.38	-0.30	0.30	0.50	1.00	
S	0.47	0.10	0.19	0.02	0.53	0.45	0.43	-0.30	0.22	0.98	0.64	1.00

IV. C. 4) Prediction of academic performance based on online learning behavior

In this section, from the data collected above, the learning behavior data and the individual final exam scores of 300 students studying a course on an online learning website in 2024 are selected. A Bayesian network-based online learning assessment system was constructed to predict the academic performance of the students.

The converged results were obtained by running the program 1000 times using Stata/MP 14.0. The estimated goodness-of-fit R^2 of the model is 0.795, indicating a good model fit. The final grades obtained from the fitting of each student were compared with their true final grades, and the results of the comparison between the model predicted results and the true results were obtained as shown in Figure 4. It can be found that the trend direction of the final grade curve obtained by the model and the real final grade curve is consistent, indicating that the grade prediction accuracy of the system is relatively high. In other words, the learning assessment model based on Bayesian network can predict students' academic performance to a certain extent according to their online learning status.

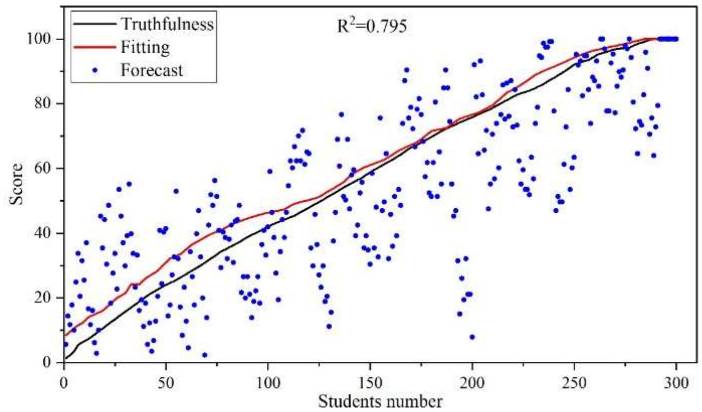

Figure 4: The comparison of the model prediction results with the real results

V. Conclusion

Through comparative experiments, the Bayesian network-based online learning assessment system proposed in this study shows high assessment and prediction accuracy, with an assessment accuracy of 92.65% and a prediction accuracy of 90.84%. Compared with other commonly used algorithms, the method proposed in this paper has obvious advantages in both assessment accuracy and prediction accuracy. In particular, during the assessment process, the Bayesian network can effectively capture the characteristics of students' learning behaviors and adjust the assessment results in real time according to the actual performance of the students, thus realizing personalized learning guidance.

Further analysis shows that the error of the system is small, with an absolute error of 2.85% and an average error of 3.68%. This result shows that the Bayesian network-based model can effectively reduce the prediction error and improve the accuracy of learning assessment. Meanwhile, in terms of data processing, the model is able to handle large-scale online learning data better, providing an efficient learning assessment solution for educational platforms. The experimental results demonstrate that the proposed assessment system not only improves the accuracy of online learning assessment, but also provides an effective decision support tool for educators and learners, which helps to further promote the implementation of personalized education.

References

[1] Lockee, B. B. (2021). Online education in the post-COVID era. Nature Electronics, 4(1), 5-6.

[2] Du, X., Zhang, M., Shelton, B. E., & Hung, J. L. (2022). Learning anytime, anywhere: a spatio-temporal analysis for online learning. Interactive Learning Environments, 30(1), 34-48.

[3] Chang, J. (2021, March). Online instruction and offline classroom teaching: A study on parallel education systems. In 2021 9th International Conference on Information and Education Technology (ICIET) (pp. 157-160). IEEE.

[4] Farley, I. A., & Burbules, N. C. (2022). Online education viewed through an equity lens: Promoting engagement and success for all learners. Review of Education, 10(3), e3367.

[5] Stoian, C. E., Fărcașiu, M. A., Dragomir, G. M., & Gherheș, V. (2022). Transition from online to face-to-face education after COVID-19: The benefits of online education from students' perspective. Sustainability, 14(19), 12812.

[6] Jindal, A., & Chahal, B. P. S. (2018). Challenges and opportunities for online education in India. Pramana research journal, 8(4), 99-106.

[7] Fung, C. Y., Su, S. I., Perry, E. J., & Garcia, M. B. (2022). Development of a socioeconomic inclusive assessment framework for online learning in higher education. In Socioeconomic inclusion during an era of online education (pp. 23-46). IGI global.

[8] Xiao, C., Qiu, H., & Cheng, S. M. (2019). Challenges and opportunities for effective assessments within a quality assurance framework for MOOCs. Journal of Hospitality, Leisure, Sport & Tourism Education, 24, 1-16.

[9] Ardiansyah, A., Al Anshori, T., Zakaria, Z., & Cahyanto, B. (2022). Principles of online learning assessment: A literature review between Western education theory and Islamic education theory. Jurnal Pendidikan Agama Islam, 19(1), 13-28.

[10] Amer, A., Sidhu, G., Bo, Z., & Srinivasan, S. (2022). A Short Review of Online Learning Assessment Strategies. International Journal of Pedagogy and Teacher Education, 6(2), 89-103.

[11] Rambe, S. L. (2021). Creative techniques for online learning assessment. Englisia: Journal of Language, Education, and Humanities, 8(2), 138-147.

[12] Ogange, B. O., Agak, J. O., Okelo, K. O., & Kiprotich, P. (2018). Student perceptions of the effectiveness of formative assessment in an online learning environment. Open Praxis, 10(1), 29-39.

[13] Elzainy, A., El Sadik, A., & Al Abdulmonem, W. (2020). Experience of e-learning and online assessment during the COVID-19 pandemic at the College of Medicine, Qassim University. Journal of Taibah University Medical Sciences, 15(6), 456-462.

[14] Usher, M., & Barak, M. (2018). Peer assessment in a project-based engineering course: comparing between on-campus and online learning environments. Assessment & Evaluation in Higher Education, 43(5), 745-759.

[15] Kondo, N., & Hatanaka, T. (2019). Modeling of learning process based on Bayesian networks. Educational technology research, 41(1), 57-67.

[16] Bowen Gu,Hao Liu,Tong Zhang,Yuan Gao,Wen Chen,Zhigang Liu... & Chonggan Liang. (2025). Hydrogen leakage risk analysis of hydrogen fuel emergency power vehicle multi-scenario based on dynamic Bayesian network. International Journal of Hydrogen Energy,136,173-185.

[17] Nanxi Wang,Min Wu & Kum Fai Yuen. (2025). Dynamic enterprise resilience assessment for port systems: A framework integrating Bayesian networks and Dempster-Shafer evidence theory. Reliability Engineering and System Safety,262,111105-111105.