Digital Education Assessment Model Based on Big Data and Its Application Under E-education

Tingzhang Yang[1], Tao Yang[2(✉)], and Xinyi Xie[3]

[1] HPI Technology Co., Ltd., Beijing, China
[2] Liuzhou Urban Rural Planning and Design Research Institute Co., Ltd., Liuzhou, China
147717112@qq.com
[3] School of Architecture and Urban Planning, Chongqing University, Chongqing, China

Abstract. Currently, our educational philosophy emphasizes the holistic development of e-learning and the promotion of educational informatization, the corresponding evaluation model has been improved from a result-based evaluation that only looks at scores and promotion rates to a process-based evaluation that values experience, participation, and thinking development. Online learning intellectual quality assessment is based on psychological and behavioral information about students' online learning to judge the learning process. The difficulty lies in the fact that there are many factors involved and there are multiple causal relationships between each factor. To address this challenge, this article proposes a process evaluation model for online learning quality using Bayesian neural networks (BNN) with coupling processing ability. The model is visualized using GeNIe software, and sensitivity analysis is performed to determine the ranking of sensitive factors that affect the quality of the online learning process. These results can assist teachers in gaining a better understanding of their students' physical and mental states, as well as their mastery of knowledge and skills during online learning. Additionally, the results can provide a foundation for educational leadership and personalized guidance.

Keywords: Smart Teaching · Educational leadership · Bayesian neural networks · Education Assessment · E-education

1 Introduction

Education informatization is a dynamic and systematic process centering on information technology, promoting teaching reform, cultivating innovative talents and building a "learning society". With the emergence of new media and technologies, such as cloud computing, big data, wireless network technology, and the Internet of Things, education informatization has become increasingly important for reflecting international competitiveness and modernizing higher education in China [1]. However, as smart education and online teaching become more widespread, the issue of online learning quality has gained attention from the community. The absence of face-to-face teaching in online

D. Kumar et al. (Eds.): IEIT 2023, AHSSEH 10, pp. 954–963, 2023.
https://doi.org/10.2991/978-94-6463-230-9_115

learning makes it difficult for teachers to monitor students' learning processes. Learning assessment as a method to promote learning, especially learning process assessment emphasizes the monitoring of the learning process and becomes an effective method to ensure the quality of online teaching. Therefore, process evaluation becomes the key to the evaluation of online learning process [2]. At present, there is more literature on the evaluation of student learning process quality, but it rarely focuses on the student learning process; moreover, the evaluation method is mainly qualitative analysis, and the quantitative analysis is also mostly scored by experts, which may lead to errors in the results; in addition, the articles on quantitative analysis of learning process quality are basically studied from the perspective of online course platforms or software developers, and the researchers are mostly from computer and software engineering backgrounds, although they adopt a more scientific evaluation method, they do not have teaching experience and are prone to ignore the cognitive laws and emotional needs of students. Process evaluation emphasizes that all activities that have educational value should be supported and affirmed by evaluation. That is, while focusing on the results, it focuses on the recording and evaluation of students' learning process and promotes the deep integration of the evaluation process and the learning process. The focus of evaluation shifts from test scores, which account for the largest proportion, to the gains in resource use, interaction and collaboration, and emotional and attitudinal aspects of the learning process. [3] The evaluation model proposed in this paper uncovers the connection between students' online learning behaviors and learning outcomes, making it easy for teachers to adjust their teaching methods. Thus making it possible to improve educational leadership [4].

2 Methodology

2.1 The Establishment of Online Learning Process Evaluation Index System

In order to determine evaluation indicators, the student learning process must first be analyzed. Gagne proposed an information processing theory that explains the learning process well. He believes that the process of student learning is similar to the work of a computer, in which the learner inputs a series of information through the brain, which is then stored, converted, processed and output by the brain to achieve interaction with the external environment, as shown in Fig. 1. The stimuli in the surrounding environment act on the receptors (sensory organs) of the learner, which transmit information to the central nervous system in the form of nerve impulses [5].

Professor Marzano, a senior professor at the National Education Laboratory, has applied the innovative findings of brain science on the learning mechanism of the brain. The study of the learning process in students illustrates that there are four major systems of thinking that are involved in the learning process, as shown in Fig. 2. The ego system helps learners determine whether they want to engage in a new task and determine their motivation to complete it. If a new task is accepted by the student, the metacognitive system is activated and the student is motivated to complete the task. If the new task is accepted by the student, the metacognitive system is activated to develop goals and plans for completing the task. Learning strategies reflect the learner's ability to adapt to the learning activity, including identifying learning methods and using learning resources

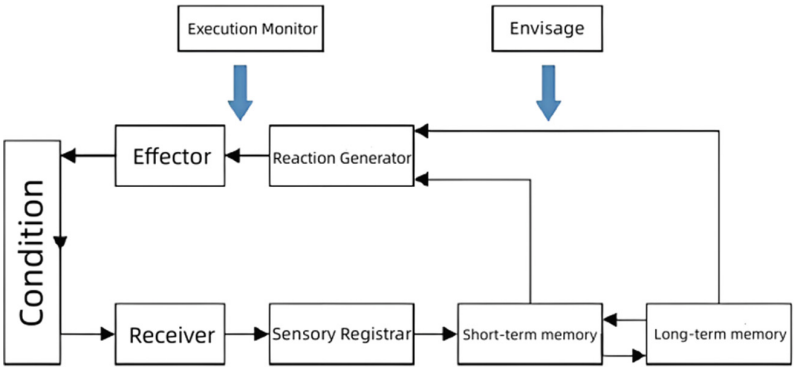

Fig. 1. Gagne Information Processing Model

effectively. The cognitive system is responsible for processing the information received to complete the task. The cognitive system is responsible for processing the information received to complete the task, which involves thinking activities such as analysis, comparison, induction, processing, and deduction. The knowledge system refers to the individual's brain's task-related knowledge and skill reserve in the individual's brain. Marzano's model of learning emphasizes the importance of student initiative, thinking development, and emotional experience as the value of the learning process. and the value of the learning process.

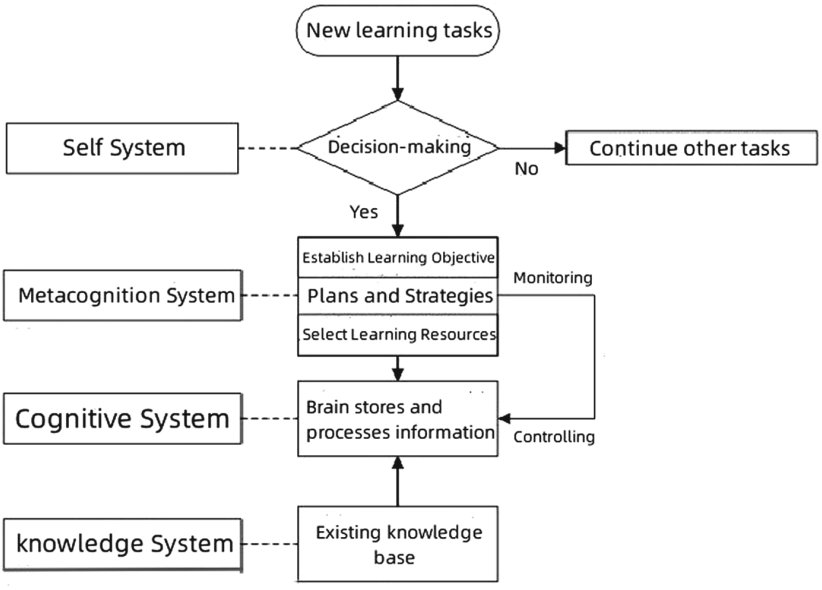

Fig. 2. Brain learning behavior patterns

Table 1. Two-level evaluation indicators and corresponding online learning behaviors

First-level indicators	Second-level indicators	Students' online learning behavior
Learning Motivation	The importance of knowledge	X1 Effective online hours, login frequency X2 Parents' scores based on student behavior observation X3 Self-rated scores
	Emotional state	
	Learning efficiency	
Learning Strategies	Objective Setting	X3 Frequency of live playback and class download X4 Frequency of speaking and asking questions in class X5 Frequency of reposting, reposting and mutual evaluation in class
	Methods'	
	Resources Utilization	
	Engagement	
	Self-Monitoring and Conditioning	
Cognitive skills	Information Acquisition	X6 Quality scoring of study notes X7 Quality scoring of mind maps X8 Formative assessment results X9 Results of group work
	Comprehension, induction	
	Knowledge application, problem solving	
Creative Thinking	Critical Thinking	X10 Reflective Summarization X11 Creative deformation question correctness
	Creative Thinking	
Information literacy	Information, resource access	X12 Information retrieval processing speed

The behaviors that make up the online learning process include logging in and browsing online teaching platform, online interaction, posting, now online learning resources, homework, etc. However, the complexity of learning behavior information makes statistical work more difficult, so it is necessary to reasonably select quantitative parameters of online learning behaviors and extract typical behaviors that best reflect a certain characteristic. According to the taxonomy of Marzano's educational objectives (Fig. 2) and Gagne's information processing model (Fig. 1), five first-level indicators are identified. In order to make a more targeted evaluation of learners' behaviors in the online learning process, the first-level indicators need to be refined and the second-level indicators outlined according to their functional connotations, as shown in Table 1.

In actual teaching, teachers should adjust the index system according to the unique characteristics of each course to meet the requirements of online teaching and process evaluation.

2.2 Bayesian Process Evaluation Model

The traditional evaluation model assumes a linear relationship between each indicator on learning quality, which is not consistent with the fact that, for example, the length of login time reflects the motivation to learn, and students with longer login time are also more

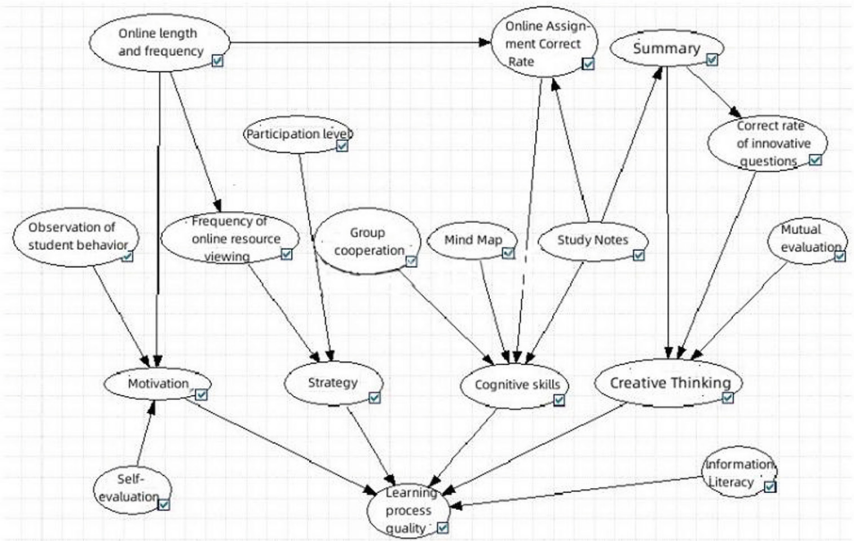

Fig. 3. Bayesian network structure chart

likely to get good grades in online assignments, so there are multiple causal relationships between each indicator. The Bayesian network offers a significant advantage in that it fully considers the causal relationships among the indicators. The network structure is constructed by determining the interrelationships, as illustrated in Fig. 4. Subsequently, the network structure is learned, and parameters are optimized using a specific algorithm to enhance the model. The conditional probability table of each node is then determined to facilitate probability calculation and reasoning. The Bayesian network model is a quantitative model that utilizes Bayes' theorem to infer uncertain probability problems, as demonstrated in Eq. (1).

$$P(A|B) = \frac{P(A|B)}{P(|B)} \tag{1}$$

Bayesian network is a graphical model describing the qualitative and quantitative dependencies between variables and is a powerful tool for joint data analysis and prediction [7], and the Bayesian model established in this paper is shown in Fig. 3

3 CASE Study and Result Analysis

3.1 Educational Data Mining

The study was conducted at a school with 524 students learning online during the online class period of the COVID-19 epidemic, with 5 lessons of 20 min each per day. Teachers uploaded 2 copies of the learning resources for students to download each day, and the online platform included modules for learning exchange, quizzes, and teacher-student evaluation.

Table 2. Raw Data of students' online behavior

Factor	Number	
	1	2
X1(time)	2	3
X2(score)	70	87
X3(score)	88	92
X4(times)	0	2
X5(times)	4	2
X6(score)	67	83
X7(score)	88	90
X8(score)	80	77
X9(score)	85	83
X10(score)	89	74
X11(Completion)	78%	75%
X12(score)	88	81

The process management module records the raw data of students' online behaviors, as shown in Table 2, which includes variables such as login time, number of discussions, etc. Therefore, it is necessary to transform the original high-dimensional dataset into a new low-dimensional dataset, e.g., if the criterion for X1 (frequency of real-time playback and classroom download) is 2 times/day and students with the number 1 meet the criterion, X3 is recorded as 1. And so on, inputting the original data into the model with 0 and 1 reduces the dimensionality of the model input variables, thus improving the efficiency of the model.

Translated with www.DeepL.com/Translator (free version).

3.2 Pre-processing

To validate the accuracy of the model, online outcome evaluations were conducted at the conclusion of all online teaching courses. These evaluations were categorized into four levels: A, B, C, and D. The comprehensive evaluations of levels A and B were recorded as 1 in X13. The X13 comprehensive evaluation results were then compared to the process evaluation results generated by the Bayesian network model (Table 3).

3.3 Calculate Conditional Probability

The quality of students' online learning process is the result of a variety of factors that arise and interact with each other at the same time. As in the previous As the index system established in the previous section, five primary indicators have a direct effect on the quality of teaching and learning, and the study notes (secondary indicator) are part of the cognitive skills (primary indicator). However, it has been observed that students who

Table 3. Pre-processed data according to evaluation criteria

Factor	Number	
	1	2
X1	1	3
X2	0	87
X3	1	92
X4	0	2
X5	1	2
X6	0	83
X7	1	90
X8	1	77
X9	1	83
X10	1	74
X11	0	75%
X12	1	81

attain the standard in study notes are more likely to achieve the standard in reflection and summarization. It is important to note that the impact of primary and secondary indicators on the quality of learning is not independent and linear. For instance, when the length of time logged in is taken into consideration, the probability of meeting the standard for online assignments is 89%, as depicted in Table 5. This probability is determined using Bayesian networks, and the conditional probability of each node variable is calculated sequentially as follows (Table 4).

Table 4. Probability table of login time

P(X1 = 0)	P(X1 = 1)
0.28	0.72

Table 5. Probability table of student motivation conditions

Condition	P(Online work) = 0	P(Online work) = 1
P(X1) = 0 P(X7) = 0	0.73	0.27
P(X1) = 0 P(X7) = 1	0.79	0.21
P(X1) = 1 P(X7) = 0	0.86	0.14
P(X1) = 1 P(X7) = 1	0.92	0.08

3.4 Analytical Calculation

GeNIe is a graphical Bayesian software developed by the Decision Applications Laboratory of the University of Pittsburgh, USA, which is widely used because of its simplicity of operation. After obtaining the conditional probability table for each node, it is input into the GeNIe software to simulate the inference on the quality of the learning process. With the results obtained from the Bayesian network structure in Fig. 3, the analysis of sensitive factors affecting the quality of the learning process can be conducted using the GeNIe software, as demonstrated in Figs. 4 and 5.

To conduct sensitivity analysis on each factor, we set the conditional probability of each parent node to 100% in turn in the Bayesian network shown in Fig. 3. We then calculate the probability change value (ΔP) of the child nodes, which represents the sensitivity of the network to each factor. To illustrate this process, we present Figs. 4

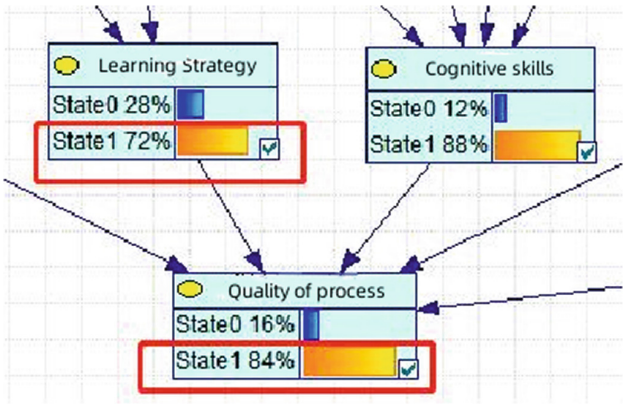

Fig. 4. Learning strategy is normal

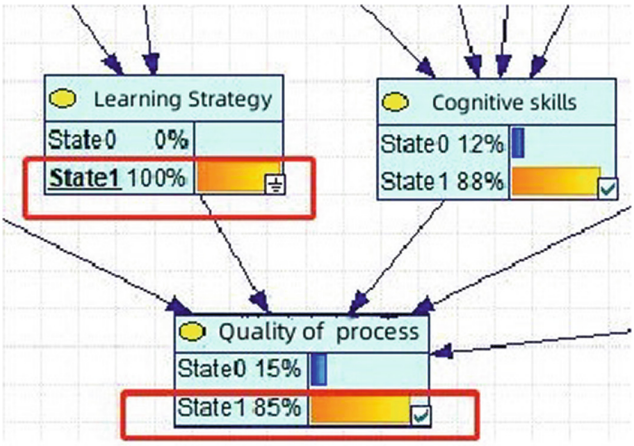

Fig. 5. Initial probability is adjusted to 100% after reaching the target

Table 6. Highest Sensitivity Ranking Teaching Quality Node

	Parent Node	Sensitivity
Teaching Quality	Learning Motivation	5%
Learning Motivation	Login frequency and average length of stay	4%
Cognitive skillsskills	Mind map	5%
Learning Strategies	Participation rate of discussion	12%
Creative Thinking	Summarize and reflect	20%

and 5, which demonstrate how we calculate the sensitivity of the quality of the learning process to the learning motivation.

From the results of the sensitivity analysis above, it can be concluded that in the Mazano taxonomy of learning goals, the most influential factor in the quality of learning is motivation. The size of students' motivation for learning plays a decisive role in the cognitive process. Students' motivation for learning comes from four main aspects: importance test, efficacy test, affective test and overall motivation test. Also, according to Table 6 the sensitivity of learning strategies on discussion participation reached 12%. The sensitivity of cognitive skills to correctness of online assignments was 4%. Currently, teachers use online assignments as an important basis for evaluating students' learning quality in online teaching, but rarely refer to the content of students' discussions and replies in forums and chat rooms. According to the results of the sensitivity analysis, students' discussion behavior has a more important impact on the quality of learning.

4 Conclusion

1) Based on the Mazano education goal classification theory, a two-level evaluation system for online learning quality was established, and the causal relationships among the indicators were determined. The data on online learning behavior of 524 students was collected and pre-processed, and the quality of their learning process was evaluated using Bayesian network. The model results were found to have an error rate of less than 5% when compared to the actual test results, which indicates the high reliability and validity of both the evaluation index system and the model.
2) Using the probability calculation and visualization software GeNIe, we ranked the sensitivity of the factors influencing online learning quality. The sensitivity of "learning quality to learning motivation" was found to be 5%, ranking first among the primary indicators. Instruction is designed to raise learners' expectations of what they can accomplish, to ensure that they activate the metacognitive systems in their brains, and to ensure continued engagement in subsequent learning. Critical and creative thinking, learning strategies, and cognitive skills ranked second to fourth, while learning strategies were found to be 12% sensitive to participation in discussions in the second tier. With a sensitivity of only 4%, cognitive skills are highly dependent on the accuracy of online assignments. Hence, teachers must take an active role in

guiding and monitoring students' interactive discussions to enhance their educational leadership in the online learning process.

References

1. Han, Kwai-Koi. The construction of science teacher literacy model in the context of educational informatization[J]. Western Quality Education,2023,9(05):141-144.
2. Shen H. On the construction of hybrid online special courses based on education informatization environment [J]. Computer Knowledge and Technology, 2022, 18(10): 159-160
3. Nixon Jon. Learning to lead for transformation: an African perspective on educational lead-ership [J]. Comparative Education,2023,59(1). HAMRANI A, AKBARZADEH A, MADRAMOOTOO C A, et al. Machinelearning for predicting greenhouse gas emissions from agricultural soils [J]. Science of the Total Environment, 2020, 741: 140338.
4. Du Rujuan. Research and implementation of process evaluation based on e-learning [D]. Northwestern University, 2016.
5. Keshmiri S, Momeni Rad A. Effect of Using the Electronic Content Designed Based on Robert Gagne Model on the Level of Student Learning in Statistics Lesson[J]. Education Strategies in Medical Sciences, 2015,8(3).
6. Y. Suryani, Sapriya, E. Malihah, K. Komalasari. Developing investigation group learning model based on Marzano instructional framework to promote students' higher order thinking skill [J]. Journal of Physics: Conference Series,2020,1477(4).
7. Xing W, Li C, Chen G, et al. Automatic Assessment of Students' Engineering Design Performance Using a Bayesian Network Model[J]. Journal of Educational Computing Research, 2020, 59(2):073563312096042.