

Construction and Validation of a Neural Network-Based Blended Teaching Quality Evaluation Model

Dongxue Tu^{1,2}

¹Department of Basic Course Teaching and Research, Guangdong Technology College, Zhaoqing, Guangdong, China

²Faculty Of Arts, Communication & Education, Centre for Postgraduate Studies (CPS), Kuala Lumpur University of Science and Technology, Selangor, Kuala Lumpur, Malaysia

Abstract: Addressing the persistent challenges in blended teaching quality assessment—overly intricate indicator systems, pronounced subjectivity, and the inadequacy of conventional evaluation approaches—this study takes mathematics courses as a case study and proposes a neural network-based evaluation framework. From offline questionnaires and online teaching platforms, 118 valid samples were obtained. Following min-max normalization, two distinct models were constructed: a standard BP network configured with a 20-7-1 architecture, and a genetic algorithm-enhanced variant. The latter utilized the evolutionary algorithm to optimize initial connection weights and thresholds, resulting in a 20-41-1 topology. MATLAB was employed for both training and testing. The experimental findings show that the standard BP model produced an average error of 4.61 and a relative error of 0.05. By contrast, the model refined through genetic optimization reduced these figures to 1.94 and 0.02, respectively, while also demonstrating a markedly narrower error spread. Its predictive robustness surpassed not only that of the basic BP model but also that of the benchmark algorithms (GA and BSA). The outcomes confirm that the hybrid GA-BP approach achieves significantly higher accuracy and reliability for blended teaching quality evaluation, offering a dependable scientific basis for instructional management decisions.

Keywords: Blended Teaching; Teaching Quality Evaluation; Genetic Algorithm Optimization; Mathematics Courses

1. Introduction

The blended teaching model, which combines

the strengths of online and traditional instruction, has been gradually applied across various disciplines. However, research on how to establish a comprehensive and effective evaluation model for blended teaching remains limited. The autonomy and flexibility inherent in blended teaching, along with the complexities of its implementation, make the development of an evaluation system particularly challenging. Traditional teaching quality evaluation systems are not suitable for the current blended teaching paradigm.

Although blended teaching research has advanced considerably worldwide, the assessment of teaching quality – a key part of blended education – frequently remains shallow or incomplete because of statistical difficulty and numerous interfering factors. Currently, no generally accepted or optimal evaluation model exists for blended teaching quality.

Based on instructional practice, this study first establishes a practice-grounded evaluation framework for blended teaching, and then takes mathematics courses as the subject to construct a BP neural network-based assessment model. To address the inherent limitations of the standard BP network, a genetic algorithm-optimized variant—referred to as the GA-BP model—is introduced for quality evaluation in blended teaching settings. Subsequently, an empirical comparison of the two approaches is conducted.

2. Collection and Preprocessing of Raw Data

In the research and practice of blended teaching, systematically evaluating teaching effectiveness often requires integrating multi-source data for analysis. Based on blended teaching theories and practices, this study identifies key factors influencing evaluation outcomes and collects data from both offline classrooms and online platforms.

Indicators related to offline classes were

obtained via custom-built questionnaires. The items followed the evaluation indicator framework, and open-ended questions were omitted from scoring to preserve analytical validity. Following distribution and retrieval, the questionnaires underwent systematic review and filtering, and the final valid responses were entered into the dataset. The required online data were directly retrieved from the UOOC and Rain Classroom platforms, whose back-end systems allowed the export of relevant information in alignment with the pre-established assessment framework.

Following the integration and cleaning of both the survey responses and the online platform records, a final dataset of 118 valid samples was compiled. To eliminate the challenges posed by differences in data dimensions and value ranges in determining weights, all scoring data are normalized and scaled to the $[0,1]$ interval. This study employs the min-max normalization method, which preserves the distribution structure and information integrity of the original data to a great extent through linear transformation. The calculation formula is as follows: [1]

$$X = \frac{I - I_{\min}}{I_{\max} - I_{\min}} \quad (1)$$

Here, X corresponds to the normalized input fed into the BP network, I denotes the original raw data, and I_{\min} and I_{\max} represent the smallest and largest values observed among the raw scores, respectively.

3. BP Evaluation Model

Backpropagation networks are characterized by a powerful capacity for nonlinear approximation, enabling them to capture both linear and nonlinear patterns embedded within high-dimensional, intricate datasets [2]. Leveraging this capability, the current study applies the BP model to the task of blended teaching quality assessment [3].

3.1 Model Structure Design

A standard BP neural network comprises an input layer [4], one or more hidden layers, and an output layer. The neuron counts per layer and the depth of the network must be tuned to the specific application. A well-chosen architecture shortens training time and improves the model's learning precision. Here, the evaluation indicators of the blended teaching system are

taken as network inputs, and the corresponding final scores as outputs. Using an adequate sample size, the network adjusts its connection weights during training to produce reliable predictions of teaching assessment outcomes [5,6]. The model was constructed according to the following steps.

(1) Input Layer Design

Drawing upon the pre-established blended teaching quality framework, a total of 20 second-level indicators are designated as the network inputs. Accordingly, the input layer is configured with $m=20$ neurons.

(2) Output Layer Design

The overall teaching quality score is adopted as the target output of the BP network. Hence, the output layer consists of a single neuron, i.e., $n=1$.

(3) Hidden Layer Design

Given the architectural properties and learning dynamics of backpropagation networks, increasing the number of hidden layers tends to escalate the training difficulty [7]. In accordance with Kolmogorov's theorem, the present work adopts a single-hidden-layer BP configuration. A three-layer BP architecture is not only capable of approximating any continuous function to an arbitrary degree of precision, but also offers structural simplicity and ease of implementation.

(4) Determination of the Number of Hidden Layer Neurons

To date, there is no theoretical guidance for determining the optimal number of hidden layer neurons.

According to the relevant literature, this value is conventionally derived from a set of empirical expressions and extensive computational experiments. (evaluating the convergence performance of the network) [8-10]. The following empirical equation was used to estimate the hidden-layer neuron count:

$$l = \sqrt{m+n} + a \quad (1 < a < 10) \quad (2)$$

Here, l corresponds to the neuron count within the hidden layer, m to the input layer size, and n to the number of nodes in the output layer. Based on this formula, the suitable hidden-layer size falls between 6 and 14. Through repeated trial experiments, it was found that the BP network achieved optimal performance with 7 hidden neurons.

(5) Selection of the Neuron Activation Function

Given the demands of the present investigation and the well-recognized merits of the sigmoid

function in classification tasks and functional approximation, it is chosen as the neuron's activation function. The analytical form is expressed as follows:

$$f(x) = \frac{1}{1 + e^x} \tag{3}$$

(6) Configuration of the Network Architecture Based on the outcomes derived from the preceding steps, the final model is configured as a three-layer backpropagation network with a 20-7-1 topology, as shown in Figure 1.

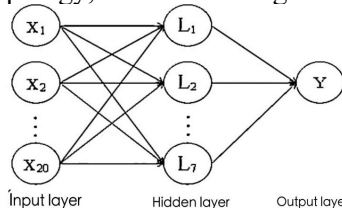


Figure 1. Structure of the Model

(7) Initialization

Prior to the network training phase, one must specify an initial range for its weights and thresholds to avoid beginning in flat gradient zones and to reduce the risk of local minima. Typically, small random values are assigned to weights, which helps shorten the training duration. For this work, the initial values of all connection weights and thresholds were confined to the interval $[-0.5, 0.5]$. Following the procedure detailed in the preceding sections, the BP-based blended teaching quality evaluation model was constructed, as illustrated in Figure 2.

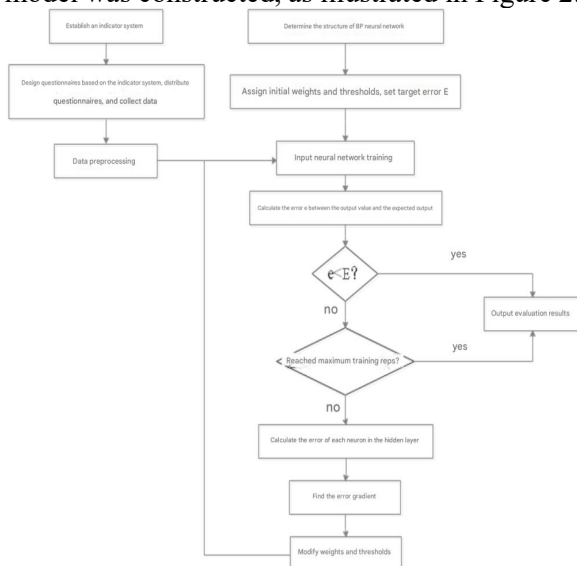


Figure 2. BP- Based Blended Teaching Quality Assessment Model

3.2 Training Sample Collection

When constructing a blended teaching quality evaluation model, the choice of samples plays a

critical role in determining both the network's learning effectiveness and the overall academic validity of the model. This research uses the "Economic Mathematics (II)" course offered in the second semester of the 2024–2025 academic year at a university in Guangdong as a case study. Following the collection and preprocessing of 118 valid data sets from the blended teaching process, 100 instances were employed for training the backpropagation network, while the other 18 were reserved as the test set to assess the resulting model's predictive accuracy and generalization ability.

3.3 Sample Training and Error Analysis

MATLAB offers strong computing power and effective graphical outputs, and it includes many toolboxes – among them the Neural Network Toolbox. This toolbox is dedicated to network system analysis and design, providing callable functions for network creation, weight initialization, training, and simulation. These functions greatly simplify the development of neural network models [11]. In this study, the BP neural network algorithm was implemented using MATLAB. The procedure comprised four main steps: The implementation of the BP neural network in MATLAB involves four main steps. First, the “init” function is called for automatic initialization. Second, the “newff” function sets the input-layer neuron count based on the sample data, and a 20-7-1 three-layer structure is adopted, with training epochs set to 5,000 and the target error to 0.0001. Third, the “train” function performs network training using the preprocessed samples. Fourth, the “sim” function evaluates the test set. Simulation experiments were carried out in the MATLAB Neural Network Toolbox, and the resulting outputs from the test set, together with the actual values, are shown in Table 1.

According to the training outcomes in Table 1, The quality prediction model built upon the BP network produced an average error of 4.61, with a corresponding relative error of 0.05. The model was then used to predict the final 18 test samples, and the corresponding prediction outcomes are illustrated in Figure 3.

As clearly shown in the figure 3, the prediction errors for the 18 sample sets are mostly within 8 points, which aligns reasonably well with actual teaching scenarios. Given that numerous influencing factors in the various stages of blended teaching cannot be fully quantified and

incorporated into the evaluation system, such errors are inevitable. On the other hand, the model can be further optimized to enhance the accuracy of the evaluation results.

Table 1. Comparison of the Two Results

Sample number	Actual Evaluation Result	BP Simulation Evaluation Result	Error	Relative Error
1	92	88	-4	0.04
2	85	81	-4	0.05
3	78	75	-3	0.04
4	96	91	-5	0.05
5	88	85	-3	0.03
6	82	78	-4	0.05
7	90	85	-5	0.06
8	83	79	-4	0.05
9	87	82	-5	0.06
10	94	89	-5	0.05
11	79	87	+8	0.08
12	91	86	-5	0.05
13	84	80	-4	0.05
14	77	84	+7	0.07
15	93	88	-5	0.05
16	81	77	-4	0.05
17	86	82	-4	0.05
18	95	91	-4	0.04

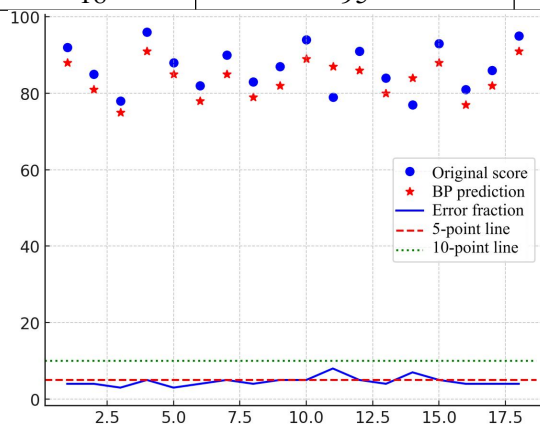


Figure 3. Prediction Results

4. GA-BP Evaluation Model

When comparing the actual evaluation outcomes with the simulated ones, it becomes evident that certain results deviate considerably. To overcome these shortcomings, this study adopts the GA-BP for quality assessment. By optimizing the BP network, a more effective evaluation model can be achieved.

4.1 BP Neural Network Based on Genetic Algorithm

The genetic algorithm is an evolutionary optimization method that simulates the collective behavior of biological populations. In this algorithm, each individual represents a candidate solution within the problem’s search space. Starting from a randomly initialized population,

the GA performs stable optimization by applying inheritance and mutation operations, thereby driving the population toward better regions of the solution space [12]. The process of optimizing a BP neural network with a genetic algorithm consists of the following three steps: Determining the BP Network Architecture: Drawing upon the predefined blended teaching evaluation system, the input dimension is set to 20 in accordance with the indicator framework. The output layer is configured with a single node, corresponding to the final evaluation score. For the hidden layer, the node count is established as 41, derived from the heuristic formula $l = 2m + 1$, where m represents the input layer’s node count.

(2) Genetic Algorithm-Based Optimization of Connection Parameters: An initial population is spawned at random, with each member carrying an encoded representation of the network’s connection weights and bias terms. A fitness function is subsequently employed to evaluate the fitness of each candidate solution, from which the best-performing individual is extracted via selection, crossover, and mutation.

(3) Evaluation with the Genetic Algorithm-Refined Network: Once this optimal chromosome is used to seed the backpropagation network, the connection parameters undergo further local fine-tuning throughout the training phase. The resultant GA-BP model delivers

superior predictive accuracy and greater computational efficiency. Adhering to this workflow, the blended teaching quality evaluation model grounded upon the GA-BP hybrid approach is established, as depicted in Figure 4.

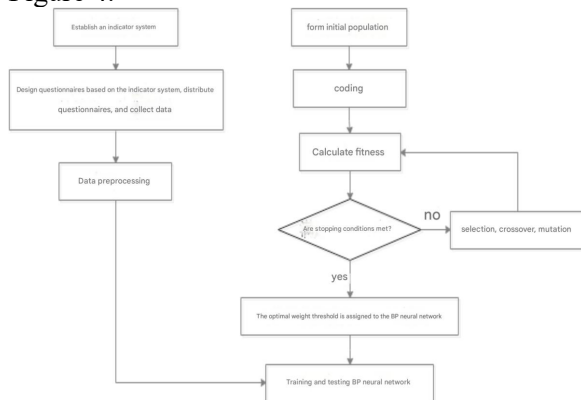


Figure 4. Evaluation Model for Blended Instruction

4.2 Model Structure Design of GA-BP Neural Network

- (1) Once the neural network topology is fixed, the preprocessed data serve as the network inputs. In this study, the GA-BP model adopts a 20-41-1 network structure.
- (2) Both the population count and the maximum number of evolutionary cycles were fixed at 2000.
- (3) The fitness value is calculated using a simplified fitness function (Equation (4)).

$$F = \frac{1}{\sum_{i=1}^N |x'_i - x_i|} \tag{4}$$

A flowchart detailing the complete GA-BP algorithmic workflow is displayed in Figure 5.

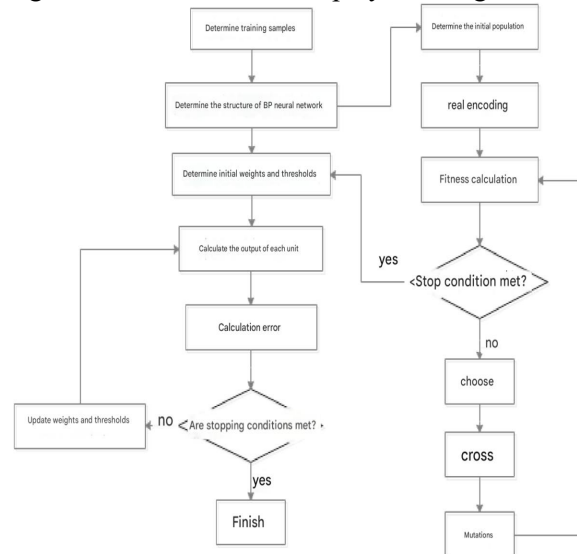


Figure 5. Flowchart

4.3 Training and Error Analysis

Using Matlab, we conducted simulation experiments to assess blended teaching quality with the GA-BP. Predictions were made for the sample data corresponding to groups 71 through 85, and Table 2 presents a comparison between the simulated evaluation results and the actual ones.

Table 2. Comparison of the Two Results

Sample number	Actual Evaluation Result	GA-BP Simulation Evaluation Result	Error	Relative Error
1	92	90	-2	0.02
2	85	84	-1	0.01
3	78	81	+3	0.03
4	96	94	-2	0.02
5	88	89	+1	0.01
6	82	81	-1	0.01
7	90	88	-2	0.02
8	83	82	-1	0.01
9	87	85	-2	0.02
10	94	92	-2	0.02
11	79	84	+5	0.05
12	91	89	-2	0.02
13	84	82	-2	0.02
14	77	76	-1	0.01
15	93	91	-2	0.02
16	81	85	+4	0.04
17	86	85	-1	0.01
18	95	94	-1	0.01

The data in Table 2 indicate that the GA-BP yields a lower overall error than the BP described earlier. The predicted values closely

match the actual ones, with an average test error of 1.94 and a relative error of 0.02. These prediction results are displayed in Figure 6.

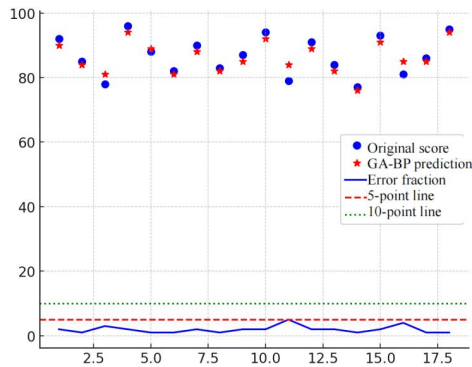


Figure 6. GA-BP Neural Network Evaluation Results

It is evident from the figure that the GA-BP’s error curve exhibits less variation relative to the actual scores than does the BP model’s. For all 18 test samples, the prediction errors remain under 5 points, with most falling within 3 points. Compared with the BP, the proposed GA-BP approach achieves greater precision and more reliable predictions.

5. Model Comparison

In this study, two models were designed, their training errors were analyzed, and blended teaching evaluation models were constructed accordingly. To further demonstrate the performance of these two neural network models, two baseline algorithms, namely GA and BSA, were also applied to forecast the same 18 test samples for comparison. The corresponding evaluation results are presented in Figures 7 and 8, respectively.

While conventional evolutionary algorithms (GA and BSA) are capable of producing predictions based on students’ various performance scores, their evaluation outcomes for the 18 datasets exhibit considerable volatility, with some predicted values even falling outside

the normal score range (exceeding 100 points). By contrast, the BP neural network—regardless of whether it is the standard or the improved version—consistently yields predictions within the valid scoring interval. Table 3 provides a comparison of the GA-BP’s predictions on the final 18 test samples against those from BP, GA, and BSA. According to the table, GA and BSA produce larger test errors than the BP, whereas the GA-BP’s predictions closely match the actual results and achieve the smallest error among the four methods. This pattern becomes even clearer when the prediction outcomes and errors are presented as comparative box plots in Figures 9 and 10.

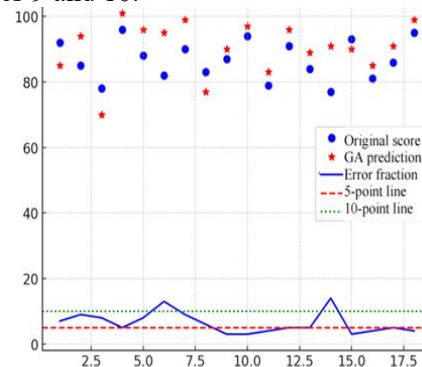


Figure 7. GA

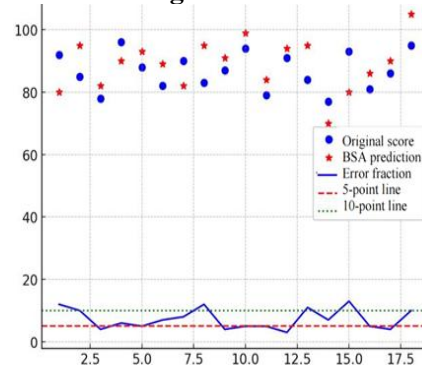


Figure 8. BSA

Table 3. Model Comparison

Sample number	GA	BSA	BP	GA-BP	Actual Evaluation Result
1	85	80	88	90	92
2	94	95	81	84	85
3	70	82	75	81	78
4	101	90	91	94	96
5	96	93	85	89	88
6	95	89	78	81	82
7	99	82	85	88	90
8	77	95	79	82	83
9	90	91	82	85	87
10	97	99	89	92	94
11	83	84	87	84	79
12	96	94	86	89	91
13	89	95	80	82	84

14	91	70	84	76	77
15	90	80	88	91	93
16	85	86	77	85	81
17	91	90	82	85	86
18	99	105	91	94	95
Error	6.39	7.27	4.61	1.94	

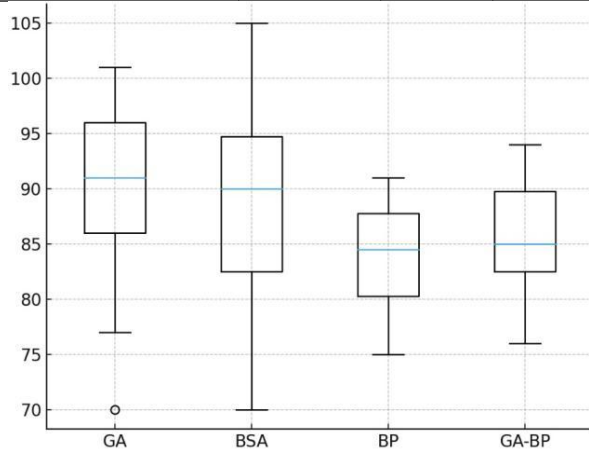
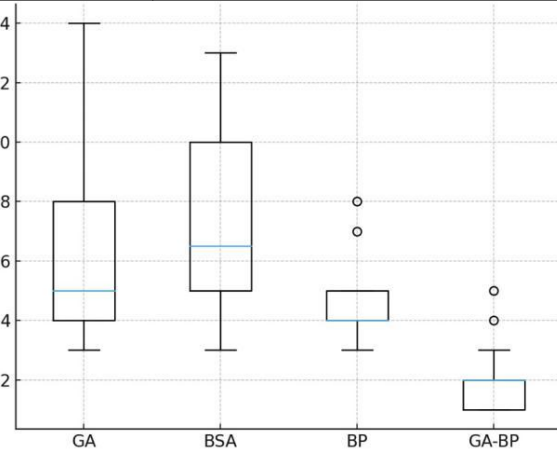


Figure 9. Box Plot of Evaluation Results Figure

Figures 9 and 10 clearly reveal distinct differences in both the evaluation outcomes and the associated errors across the four models. The original BP network produces results with considerable variability, as reflected in its box plot by a wide spread and several outliers. Although GA and BSA yield somewhat more stable predictions than the BP network, they still show notable extreme values and relatively large error ranges. By comparison, the GA-BP model's results are far more tightly clustered. Its error box plot indicates that errors generally lie within a narrow interval, with the maximum absolute error kept to about 5 points. These findings confirm that the GA-BP achieves superior accuracy, thereby offering a more stable and reliable assessment of blended teaching quality.

6. Conclusion

This research provides a reliable calculation tool for the precise assessment of teaching quality. Given the multidimensional and nonlinear nature of real-world teaching evaluation data, the work notably improved model accuracy and robustness by applying min-max normalization to original indicators and performing global optimization of the BP network's architecture and connection weights via a genetic algorithm. Experimental training and testing results indicate that the GA-BP model surpasses conventional BP, GA, and BSA models in both error control and generalization, thereby supporting a more scientific and rational evaluation of blended



10. Box Plot of Errors

teaching quality. These outcomes validate the hybrid model's effectiveness and superiority for teaching quality assessment. Furthermore, the proposed model can be integrated into practical teaching management systems, furnishing schools with strong support for process-oriented evaluation and data-driven instructional decision-making.

Acknowledgments

This paper was supported by the following projects:

Research Project of the 2024 Annual Project by the Professional Committee on Teaching Quality Management of Private Universities under the Guangdong Higher Education Teaching Management Association: Blended Teaching Quality Evaluation for Mathematics Courses Based on Neural Networks (No. GDZLGL2423)
 Guangdong Technology College 2024 "Quality Engineering" Project: Blended Teaching Quality Evaluation for Mathematics Courses Based on Neural Networks (Project No. JXGG2024150)
 Guangdong Province Higher Education Association "14th Five-Year Plan" Research Project for 2024 on the Evaluation of Mathematical Modeling Capabilities of Undergraduates in New Era Applied Universities (No. 24GQN35)
 Guangdong Technology College 2025 first-class course construction project Economic Mathematics (II) (No. YLKC202521)
 Guangdong Technology College 2025 Annual

Smart Course Construction Project: Economic Mathematics (II) (No. ZHKC202575)

References

- [1] Mei L. Model Construction of higher education quality assurance system based on fuzzy neural network. *Informatica*, 2024, 48(10).
- [2] Nan Z. Research on the application of BP algorithm of neural network in teaching quality evaluation //2024 International Conference on Optimization Computing and Wireless Communication (ICOCWC). *IEEE*, 2024: 1-4.
- [3] Huang Lu, Ren Hang, Liu Mingxi, et al. Research on the Measurement Method of Industry-University-Research Knowledge Flow Integrating Large Language Models and Dynamic Graph Neural Networks. *Chinese Journal of Management Science*, 2025, 1-16.
- [4] Li Ping, Wang Zhi, Pan Haijing. Construction of a Diversified Teaching Evaluation System for University Mathematics Courses Based on Learning Community. *Forest Region Teaching*, 2025, (10):97-100.
- [5] Zhang Hongyi, Tang Siwei. The Structure and Cultivation of Rural Teachers' Place-based Teaching Evaluation Competence. *Journal of Teaching and Management*, 2026, 1-5.
- [6] Yang H, Fu Y F. Construction of blended teaching quality evaluation indicator system for vocational colleges in the new era. *Vocational and Technical Education*, 2021, 42(35): 67-72.
- [7] Duan Wenwen. New Quality Productive Forces Empowering High-Quality Development of University Education and Teaching Management: A Review of 'Research on the Current Situation and Innovative Development of University Education and Teaching Management'. *Higher Education Exploration*, 2026, (03):2.
- [8] Xu Anfeng, Li Bin, Li Yue. Research on Online and Offline Blended Smart Education Model. *Theory and Practice of Education*, 2024, 44(15):57-60.
- [9] Jiang Hongling, Chen Linan. Construction and Practice of Online and Offline Blended Teaching Evaluation System: Taking Computer Network Course Teaching as an Example. *China Modern Educational Equipment*, 2026, (07):140-143+147.
- [10] Zhang Jingjing, Huang Yicheng. The Policy Implementation Logic of Artificial Intelligence Empowering Blended Teaching Reform. *Ethnic Education Research*, 2025, 36(06):31-41.
- [11] Lai Jun, Zhang Jinrui. Simple Implementation of Adaptive Fast Multipole Algorithm Based on MATLAB. *Mathematica Numerica Sinica*, 2025, 47(01):1-20.
- [12] Alhijawi B, Awajan A. Genetic algorithms: Theory, genetic operators, solutions, and applications. *Evolutionary Intelligence*, 2024, 17(3): 1245-1256.