*Sensors and Materials*, Vol. 37, No. 7 (2025) 2757–2777
MYU Tokyo


S & M 4085

# Metaheuristic-based Deep Learning Application for Higher Education Teaching Quality Assessment

Xin Guo,[1,2] Ying Chen,[2] Zilong Yin,[3]
Ruoying Wang,[2] Dazhi Li,[4*] and Shih-Pang Tseng[2,5**]

[1]Department of Artificial Intelligence, Shanghai Pudong Vocational and Technical College, Shanghai 201315, China
[2]School of Information Science and Technology, Sanda University, Shanghai 201209, China
[3]Department of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China
[4]The College of Information, Mechanical and Electoral Engineering, Shanghai Normal University, Shanghai 200234, China
[5]School of Software and Big Data, Changzhou College of Information Technology, Jiangsu 213164, China

(Received June 3, 2024; accepted May 26, 2025)

***Keywords:*** teaching quality assessment, analytic hierarchy process, estimation distribution of algorithm

To ensure teaching quality, higher education institutions worldwide use a teaching quality assessment (TQA) system to collect students' feedback regarding courses. How to use a more intelligent method to effectively analyze the feedback to determine practical teaching effectiveness is still the main research challenge at this stage. One of the important issues is applying sensor technology to TQA. In this study, we developed and implemented a TQA system based on analytic hierarchy process–estimation of the distribution algorithm–backpropagation (AHP–EDA–BP) to illustrate the deeper meaning of students' feedback. AHP–EDA–BP ensured the accuracy of the quantitative analysis. The AHP aggregated the opinions of experts to adjust the weights of the assessment index. The EDA–BP neural network was used to evaluate the grade of teaching quality. The experimental results showed the essential effectiveness of the proposed method. In addition, the TQA system has been applied to the TQA process at Sanda University, Shanghai, China.

## 1. Introduction

Learning is the human activity of obtaining new knowledge, skills, and abilities to better adapt to natural and social environments. On the other hand, from a social viewpoint, the education system should provide various learning services to members of society. However, the learning resources, such as qualified teachers, are usually limited and cannot meet the total learning requirement. It is an important endeavor to make use of limited learning resources more efficiently and effectively.(1) Furthermore, the education system has been an important factor in the human economic and social development of every modern country, especially the higher education system.

*Corresponding author: e-mail: lijunzhi@shnu.edu.cn
**Corresponding author: e-mail: tsengshihpang@czcit.edu.cn
https://doi.org/10.18494/SAM5173

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

The quality of learning services provided by the higher education system greatly affects economic and social development. On the other hand, the higher education system should actively respond to the requirements of socioeconomic development and government policies. Therefore, how to effectively assess the quality of learning in higher education institutions has become a very important research topic. Teaching quality assessment (TQA) is used to measure learning quality by collecting and analyzing feedback from students for one specific course. The TQA focuses particularly on a teacher's abilities and behaviors because the teacher is the planner, executive, and manager of the course. Teachers can improve their teaching abilities and skills by referring to the TQA results. Moreover, the TQA results are considered by the heads of faculties, departments, and educational programs when adjusting their course curricula and teachers' training.(2–4) TQA is based on one acceptable and realizable assumption that better teaching from the teacher can provide a better learning effect to the students.(5) TQA can provide the teachers feedback from students to improve their teaching, as shown in Fig. 1. TQA focuses on good teaching and has become an important element in the education system,(6) especially the higher education system.(7) Nevertheless, the higher education system is different in various countries and regions. However, most higher education institutions provide qualified learning courses and programs to their students. In the 21st century, TQA work has been indispensable to higher education institutions.

In recent decades, computer and communication technologies have dramatically developed and changed various aspects of modern human life, including teaching and learning. To improve the efficiency of TQA work, the Internet questionnaire has replaced the traditional paper–pencil questionnaires. The TQA system makes not only the questionnaire more efficient but also the processing and analysis more flexible and complex to illustrate the teaching quality.

To improve the quality of courses and programs, the TQA system is used to realize the effectiveness of teaching "services" by collecting feedback from students in many higher education institutions.(3) The TQA system can be an independent system or a subsystem of the E-learning platform. Various TQA systems are designed to support the different concepts of

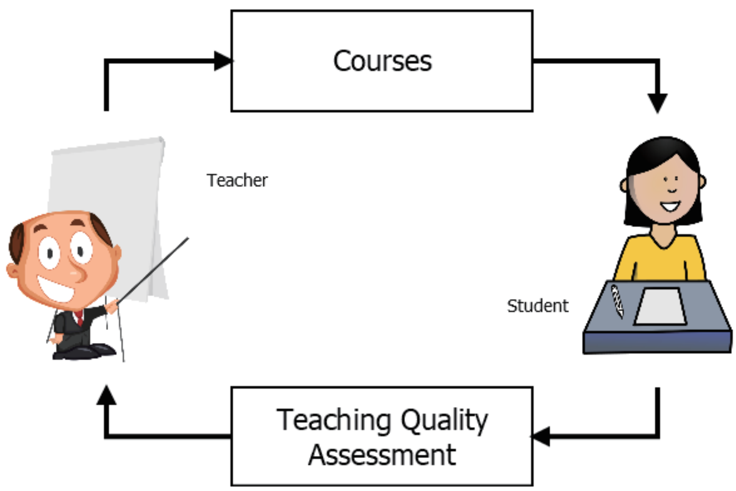



Fig. 1. (Color online) Relationship between students and teachers in TQA.

education.[8,9] In general, the result of TQA only represents the students' short-term perceived teaching satisfaction. The main research challenge at this time is how to correctly use the TQA to indicate teaching effectiveness and true teaching quality.

There are two fundamental approaches to improving the TQA system performance. The first is measuring the teaching quality more precisely. For this, one method is to modify the feedback model and structure of the students.[10,11] The second approach is to use a more intelligent analytic process to illustrate the deeper meaning of the feedback from the students.[12,13] Of course, these two approaches can support and complement each other.

Because of the fuzziness and subjectivity of TQA, it is extremely difficult to quantitatively analyze TQA directly. The analytic hierarchy process (AHP) is a hierarchical, structured, qualitative, and quantitative analysis method that is often used to calculate the weight of the teaching quality evaluation index.[9] Backpropagation (BP) neural networks have strong nonlinear mapping and self-learning capabilities and can accurately and efficiently calculate the evaluation results of teaching quality.[14] Metaheuristics[15] can optimize the parameters of the BP neural networks and improve the convergence speed and accuracy of the evaluation model. [16] In this study, the estimation of the distribution algorithm (EDA) is chosen together with the BP neural networks because the EDA has a more solid theoretical foundation. Therefore, we develop a teaching quality evaluation and analysis system based on AHP, BP neural network, and EDA.

Our novel TQA system based on the AHP and EDA–BP neural network (AHP–EDA–BP) is proposed, implemented, and applied to higher education institutions. The system has the functions of students' online teaching assessment, real-time calculation of teaching evaluation results, and data analysis. It is composed of four user authority modules: administrator module, supervisor module, teacher module, and student module. The system is developed using the Python language, Django MVC framework, and HTML5 standard. The evaluation algorithm includes AHP–EDA–BP so that the system can automatically calculate the teaching evaluation results. The results of testing and trial have proven that the system is easy to use: users can access and use it through any browser that conforms to the HTML5 standard, and the user experience is good.

The contributions of this work include the following.

- TQA has become a key task of universities. Effective methods are needed to ensure that the evaluation activities are feasible and efficient with high reliability of TQA results. This proposed TQA system can automate the TQA process and transform personal subjective fuzzy scoring into quantitative analysis to ensure the reliability of TQA results.
- Currently, the results of the TQA process are still influenced by personal subjectivity. In this work, AHP–EDA–BP is used to reduce the subjectivity of the TQA process and to increase the objectivity.
- Different experts assign different weights to the assessment indexes. The AHP is effectively used to aggregate the weights from the various experts to be more objective.
- The proposed TQA system has been implemented and applied in the physical TQA process at Sanda University. The efficiency and effectiveness of the proposed approach have been shown.

- We have solved the key technical problems in the development of the TQA system by standardizing the TQA process. As a result, decision-makers at the university can realize teaching quality in practice.

In this paper, we focus on analyzing the data collected from students to assess teaching quality. A metaheuristic-based deep learning approach is proposed to summarize the objective and subjective data from various data sources. The results unequivocally demonstrate the effectiveness and availability of the proposed approach, instilling confidence in its potential application. The teaching quality data can generally be obtained via the video/image/audio sensors in the classroom. It is of utmost importance to address the research issue of summarizing the sensing data from various heterogeneous sources for education applications, given its potential impact on the field.

The remainder of the paper is structured as follows. In Sect. 2, the related works of this research are reviewed. Brief descriptions of AHP–EDA–BP and the prototype system are given in Sect. 3. An experiment to evaluate the effectiveness of the proposed system is described in Sect. 4. The experimental results and the implementation of the prototype system are presented in Sect. 5, and the conclusions are given in Sect. 6.

## 2. Related Works

To improve teaching quality, higher education institutions popularly carry out TQA, also known as teaching quality evaluation, to realize effective teaching "services".(3,17) Collecting student feedback to evaluate the teaching quality has become the most intuitive and simple assessment method because the students are the primary consumers of teaching services. Students, as the primary recipients of teaching activity, should be the primary assessors in TQA.(18) The general TQA processes in higher education are carried out to collect students' feedback about evaluations of teaching activities on special courses, also known as course/teaching evaluations.(19,20)

In general, the TQA results from students after the courses only represent the students' short-term sense and satisfaction. However, the evolution of higher education should balance consumerism and professionalism.(21) Furthermore, higher education must still take social responsibility(22,23) and endeavor to satisfy the expectations of the whole society. Consequently, we should take account of other measurement methods in TQA.

Peer review of teaching(24–27) provides an alternative measurement method to TQA. Peer review of teaching is an integral and underused component in measuring teaching quality. Peer review of teaching can induce quality enhancement and be a powerful means of encouraging the continuing professional development of individual lecturers. In addition, the peer review of teaching can be performed by the administrators and non-administrators in higher education institutions.(26) These two groups can provide different viewpoints about teaching quality. We design a TQA system to support and integrate the feedback from students and administrator and non-administrator peer reviews.

Some computational analytic algorithms were used to evaluate the teaching quality and to elucidate the deeper meaning from the data. An AI approach proposed by Li and Su(12) focuses on the quality of teaching in online courses in elementary education. The entropy weight method and grey clustering analysis were introduced to process the data generated in the online teaching activities. In addition, several strategies were proposed to improve the quality assessment of online teaching in elementary education. Hu and Zhen[28] proposed the genetic-algorithm-based support vector machine (GA-SVM) to analyze the TQA data and score the teaching quality.

The AHP, first proposed by Saaty,[29] is a structured technique of quantitative and qualitative analysis for complex multi-objective decisions based on mathematics and psychology. In the AHP, the decision problem would first be decomposed into a hierarchy of subproblems that can be realized easily and analyzed independently. Each subproblem, an element of the hierarchy, is related to one aspect of the decision problem. In the next step, the decision-makers systematically evaluate various elements. These evaluations would be converted to numerical values that can be processed and compared over the entire range of the problem.[30] A numerical weight or priority is derived for each element of the hierarchy, allowing diverse and often incommensurable elements to be compared with one another rationally and consistently. This capability distinguishes the AHP from other decision-making techniques. In the final step of the process, numerical priorities are calculated for each of the decision alternatives. These numbers represent the relative capability of the decision alternatives to achieve the decision goal, so they allow a straightforward consideration of the various courses of action. Satty proposed five successful applications of AHP in transportation.[31] Kurttila *et al.*[32] used the AHP to enhance the strengths, weaknesses, opportunities, and threats (SWOT) analysis in a forest-certification case.

Artificial neural networks (ANNs), usually referred to as neural networks (NNs), are a category of computation frameworks based on the simulation of the natural biological neural system.[33,34] In fact, NNs are based on the linear and nonlinear combination of mathematics regressions. The combination is a multilayer structure. With the development of semiconductor technology, increasingly complex NN architectures are being proposed and constructed, consuming increasingly higher computing capacity. Modern NNs of the recent decade generally have more layers, which are called depth. Therefore, these modern NNs are often referred to as deep neural networks (DNNs).[35] For each physical NN model, the training process adjusts the parameters in accordance with the output errors from the labeled dataset. The training process is also called learning. The most popular training process uses the BP[36] algorithm. The term “backpropagation” means that the errors are propagated backward from the output layer to the input layer. The mathematical basis of BP is the application of the derivative chain rule.[37] At every node in each layer, the gradient-decent method[38] is used to adjust the parameters to reduce the error. However, there are still some shortcomings in the BP training process. Apparently, the gradient-decent method is a kind of greedy policy. Consequently, the BP process cannot efficaciously explore the search space of the parameters. To avoid the premature convergence of the BP training process, the metaheuristic algorithm is introduced as an alternative to the gradient-decent method.

The metaheuristic algorithm is a stochastic search algorithm designed to find a sufficiently optimal solution in a realistic computation time by effectively and efficiently using problem-independent and problem-dependent heuristics.[15,39] There have been various metaheuristic algorithms proposed in recent decades, such as simulating annealing (SA),[40] genetic algorithm

(GA),[41] and particle swarm optimization (PSO).[42] These modern metaheuristic algorithms can effectively balance exploration and exploitation in the search process to avoid premature convergence. However, these metaheuristic algorithms are mostly based on empirical principles and have no solid theoretical basis.

Estimation of distribution algorithm (EDA) is one category of metaheuristic algorithms based on a probability model.[43,44] Sometimes, it is also called a probabilistic model-building genetic algorithm (PMBGA).[45] EDA has a more solid theoretical basis than other metaheuristic algorithms. The traditional crossover and mutation operations of GA are replaced by the probabilistic model building and sampling of search space. Initially, the probabilistic model is assumed to be a uniform distribution. In each iteration, the probabilistic model would become more accurate as a result of the sampling operation and model-building operations. Finally, we can find the best solution by sampling the probabilistic model. In general, EDA can more effectively avoid premature convergence than the traditional GA. The probabilistic model ensures diversity in the search process.

## 3. System Design

The workflow of the TQA system is shown in Fig. 2. There are three stages in the TQA process. First, the front-end web system collects all the data from experts and students. Next, the AHP stage systematically aggregates the opinions of experts to acquire the weights of assessment indexes. Lastly, the EDA–BP network is used to evaluate the grade of teaching quality for each course.

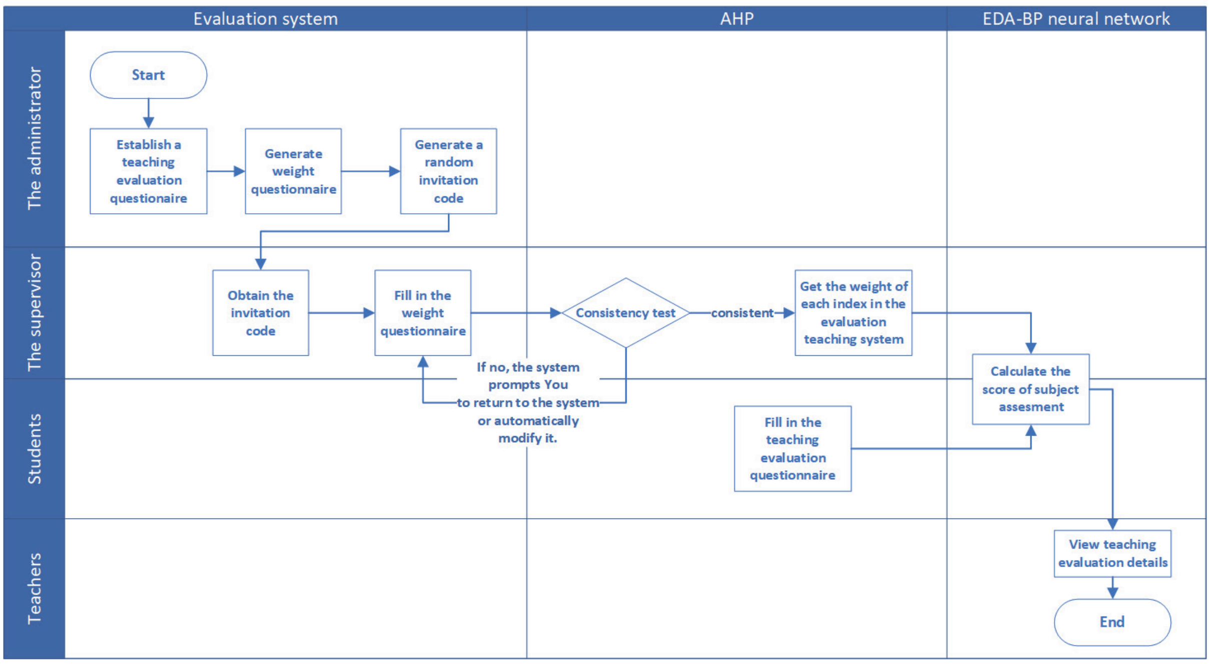



Fig. 2. (Color online) Flowchart of the system operation process, starting with the administrator creating a teaching evaluation questionnaire for teachers and ending with the teacher reviewing the evaluation result details.

In the system, the administrator first creates a TQA process. Then, the system will automatically generate a weighted questionnaire based on the TQA system, accompanied by a set of random invitation codes. The selected experts must input the invitation code from the administrator login to the system to fill out the weighted questionnaire for experts. The AHP process calculates and judges the expert's responses. By comparing the scoring assessment indexes from various experts, the average weight of each indicator in the TQA system is constructed. The average weight of each indicator is used as the final weight of the assessment indexes.

After the course is completed, the students can enter the system to score the teaching activities. The scoring data is combined with the index weights of the TQA system to evaluate the teacher's final score for teaching work. Later the teachers can log in to the system to view their TQA scores.

Because the system needs to interact with users having different needs, the functions of the system will be designed and developed in accordance with the user roles. Users are divided into four types according to their roles: administrator, supervisor, teacher, and student. In accordance with the actual business scenarios, administrators are divided into school administrators and department administrators. The functional modules are shown in Fig. 3.

### 3.1. AHP

The AHP fuzzy comprehensive evaluation method (AHP–FCEM) can essentially be regarded as a combination of the AHP and FCEM. The steps of AHP–FCEM include determining the

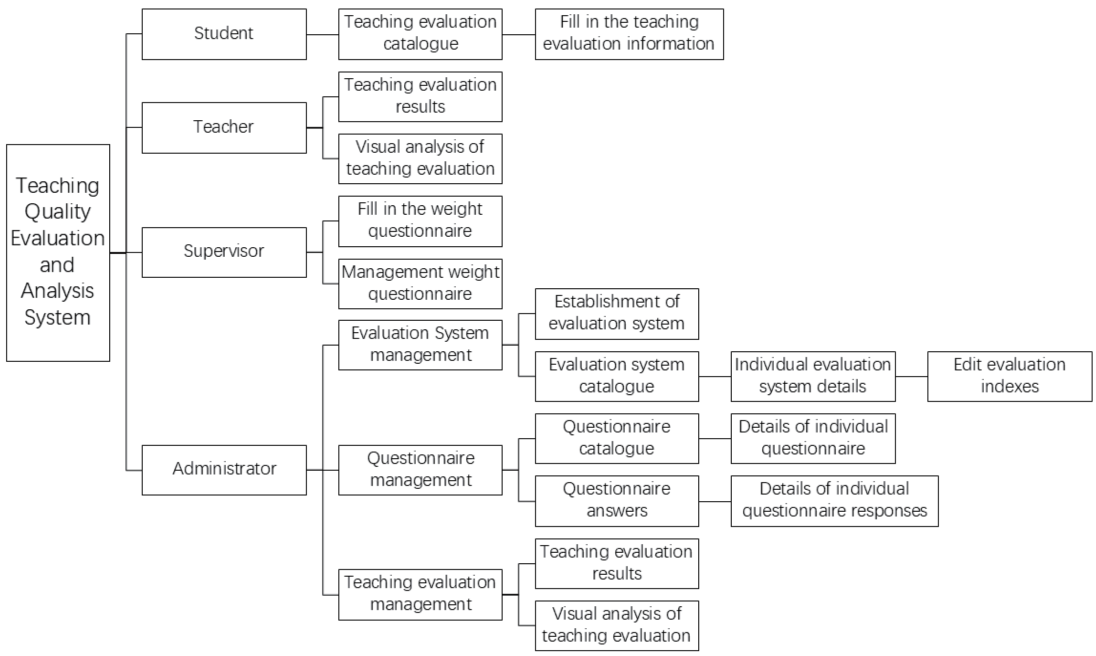



Fig. 3. System framework diagram showing the page structure and functionalities for four different users in accordance with their roles—students, teachers, supervisors, and administrators—after logging into the system.

evaluation index system, establishing the evaluation set, calculating the evaluation index weight, establishing the comprehensive evaluation matrix, and calculating the comprehensive score. In our design, the teaching content, teaching attitude, teaching skills, teaching effects, and teaching methods are the five assessment indexes used to evaluate the teaching quality. Each assessment index has five grades: excellent, good, medium, qualified, and unqualified. The corresponding scores are set to 95, 85, 75, 60, and 55, respectively. Therefore, the index set is {*Excellent, Good, Average, Fair, Poor*}, and its corresponding score set is {95, 85, 75, 60, 55}. On the basis of the questionnaire responses from the teaching experts, we construct the relative importance $n$-order matrix $\boldsymbol{W}$.

$$\boldsymbol{W} = \begin{bmatrix} w_{1,1} & \cdots & w_{1,n} \\ \vdots & \ddots & \vdots \\ w_{n,1} & \cdots & w_{n,n} \end{bmatrix} \tag{1}$$

$w_{i,j}$ means that the $i$-th index is of relative importance to the $j$-th index. Thus, $i = j \rightarrow w_{i,j} = 1$ and $w_{i,j} = 1/w_{j,i}$.

$$\bar{\boldsymbol{V}}_i = \sqrt[n]{\prod_{j=1}^{n} w_{i,j}} \tag{2}$$

We normalize the vector $\bar{\boldsymbol{V}}$.

$$\boldsymbol{V}_i = \frac{\bar{V}_i}{\sum_{i=1}^{n} \bar{V}_i} \tag{3}$$

We calculate the maximum eigenvalue $\lambda_{max}$.

$$\lambda_{max} = \sum_{i=1}^{n} \frac{(WV)_i}{nV_i} \tag{4}$$

The consistency index ($CI$) can be calculated as

$$Consistency\ Index(CI) = \frac{\lambda_{max} - n}{n-1}. \tag{5}$$

We can evaluate the consistency ratio ($CR$) as

$$CR = \frac{CI}{RI}, \tag{6}$$

where the random index (*RI*) is used to evaluate *CI*; the value of *RI* is only dependent on the dimensions of the matrix. If the *CR* is less than 0.1, the index matrix can be accepted. Otherwise, the index matrix cannot be accepted. In general, repeating the data collection is necessary to reconstruct the index matrix.

### 3.2. EDA

In AHP–EDA–BP, the results of AHP are the initial parameter values of the BP NN, and the EDA is used to adjust the parameters of the BP NN. The solution is continuously coded and represented as a floating-point vector, as shown in Fig. 4. The length of the solution vector, *L*, is equivalent to the number of parameters in the BP NN. For each parameter, the various values in the same position of different solutions can represent the result of probing this dimension in the search space. The probability distribution of this parameter can be estimated from the values in different solutions. In general, the probability distribution is assumed to be a normal distribution. After the selection operation, the set of solutions contains better solutions. Therefore, the estimation of a probability distribution can more precisely illustrate the possible region with a higher probability of finding better solutions at this stage. Thus, constructing new solutions by sampling the estimated probability distribution can help yield better solutions that can be merged into the population. After a new selection, we can estimate the probability distribution more precisely. Figure 5 shows the evolution of the probability distribution in EDA. The standard deviation is reduced because of the iteration of EDA. On the other hand, the region far from the mean still has enough probability to avoid premature convergence.

Algorithm 1 shows the general outline of EDA. First, the initial probabilistic model is assumed to be a uniform distribution. All the solutions in the population are initiated randomly because of the uniform distribution. The main loop is terminated after the predefined number of iterations. In the main loop, there are three steps: selection, building the probabilistic model, and sampling the probabilistic model. In the selection, better solutions are chosen to construct a more precise probability distribution. In the step of building the probabilistic model, the means and standard deviations of each parameter are calculated. In the step of sampling probabilistic models, new solutions are generated by sampling the distribution to determine the value for each

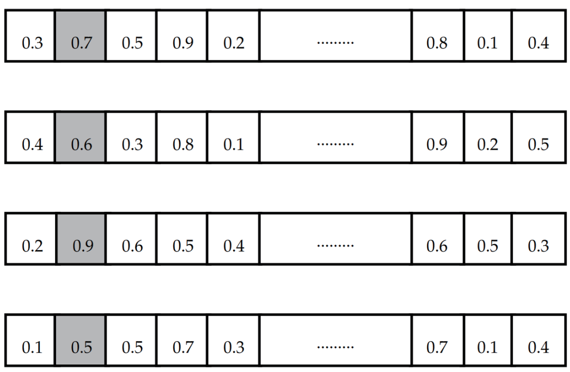



Fig. 4. Representation of solutions in EDA.

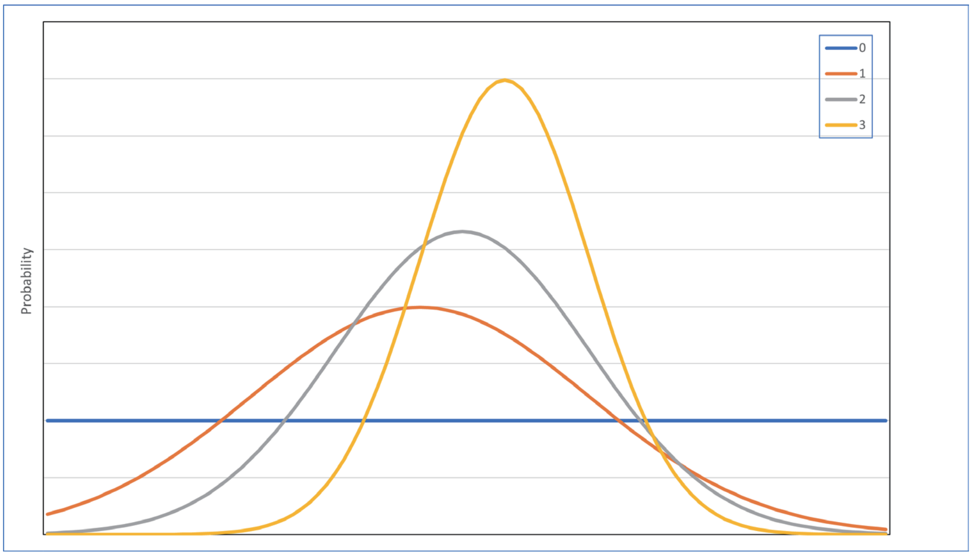



Fig. 5. (Color online) Evolution of probability distribution in EDA.

**Algorithm 1:** Estimation of Distribution Algorithm

**Input:** a set of variables and the objective function
**Output:** a set of optimized variable values for the objective function

```
generate initial probabilistic model M(0);
generate initial population P(0) by sampling M(0);
while the terminate condition is not met do
    select population of better solutions S(t) from P(t);
    build probabilistic model M(t) for S(t);
    sample M(t) to generate new candidate solutions O(t);
    add O(t) into P(t);
end while
Output the result
```

parameter. The sampling method can be directly implemented by using the normal distribution random number generator.

The compact genetic algorithm (CGA)[46] is one variant of EDA. The basic concept of CGA is to use the probability distribution to replace the population of solutions. In each iteration, there are two solutions generated by sampling the probability distribution. The binary tournament selects the winner and loser. The probability distribution is updated using the winner and loser. The parameter value of the winner solution increases its probability. Of course, the parameter value of the loser decreases its probability.

## 4. Results of Experiments

### 4.1. Indexes of assessment

In this work, the assessment indexes of teaching quality are organized as a hierarchical structure, as shown in Fig. 6. The details of each index are described in Table 1. In accordance

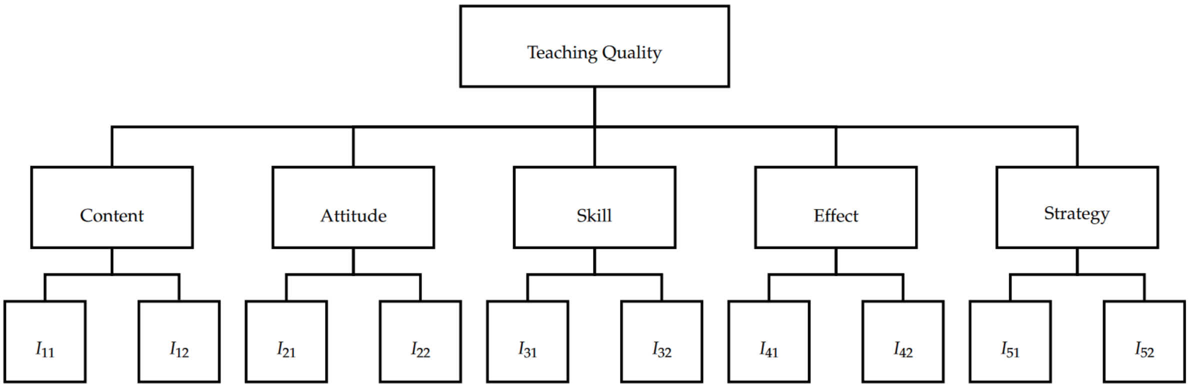



Fig. 6. The teaching quality indicator system is divided into three layers. The top layer is the target layer for teaching quality evaluation, the second layer contains five aspects for evaluating teaching quality, and the third layer has evaluation indicators refined for each of the five aspects in the second layer.

Table 1
Indexes of assessment.

| $L_0$ | $L_1$ | $L_2$ |
|---|---|---|
| Teaching quality | $I_1$: Content | $I_{11}$: Precise concepts, clear points and appropriate examples |
| | | $I_{12}$: Combined theoretical and practical frontier development |
| | $I_2$: Attitude | $I_{21}$: Well prepared, corrects exercises and answers questions |
| | | $I_{22}$: Appropriate clothing, manner and punctuality |
| | $I_3$: Skill | $I_{31}$: Modernized presentation and demonstration |
| | | $I_{32}$: Precise and standard Chinese speaking and writing |
| | $I_4$: Effect | $I_{41}$: Moderate difficulty and easy to understand |
| | | $I_{42}$: High acceptance and interaction from students |
| | $I_5$: Strategy | $I_{51}$: Flexible, multiform and innovative |
| | | $I_{52}$: Inspires the students to think and learn |

with the index hierarchy, we designed the questionnaires to obtain the weights from the experts. We use a 5-point Likert Scale to measure the degree of agreement for each index. Each response has a corresponding value to measure the degree of agreement of the response. The response set is $C$ = {*Excellent, Good, Average, Fair, Poor*}.

### 4.2. Assign weights of indexes via AHP method

The weights of all assessment indexes are assigned referring to the expert survey. More than ten teaching experts from the front line of teaching and administration positions are invited to fill out the questionnaire about the importance of each index. A consistency check is used to select ten responses to the survey questionnaire. If a questionnaire cannot pass the consistency test, it is not an effective questionnaire. We also recruit an alternative expert to fill out the questionnaire. Finally, we obtain ten effective questionnaires.

For example, in one questionnaire, the following shows the weight calculation process for $L_1$ indexes, $\{I_1, I_2, I_3, I_4, I_5\}$. On the basis of the relative importance of the five indexes from the selected experts, we can construct the following relative importance 5 × 5 matrix, $\boldsymbol{W}$.

$$\boldsymbol{W}=\begin{bmatrix} 1 & 1 & 1 & 6 & 1 \\ 1 & 1 & 0.2 & 6 & 1 \\ 1 & 5 & 1 & 7 & 1 \\ 0.167 & 0.167 & 0.143 & 1 & 0.167 \\ 1 & 1 & 1 & 6 & 1 \end{bmatrix} \tag{7}$$

In the next step, we calculate the weight of each index using

$$\boldsymbol{M}=\begin{bmatrix} \sqrt[5]{1\times1\times1\times6\times1} \\ \sqrt[5]{1\times1\times0.2\times6\times1} \\ \sqrt[5]{1\times5\times1\times7\times1} \\ \sqrt[5]{0.167\times0.167\times0.143\times1\times0.167} \\ \sqrt[5]{1\times1\times1\times6\times1} \end{bmatrix}=\begin{bmatrix} 1.431 \\ 1.037 \\ 2.036 \\ 0.231 \\ 1.431 \end{bmatrix}, \tag{8}$$

$$\boldsymbol{V}=\begin{bmatrix} 1.431 \\ 1.037 \\ 2.036 \\ 0.231 \\ 1.431 \end{bmatrix}\div\left(1.431+1.037+2.036+0.231+1.431\right)=\begin{bmatrix} 0.2321 \\ 0.1682 \\ 0.3302 \\ 0.0375 \\ 0.2321 \end{bmatrix}. \tag{9}$$

The consistency test is performed, and we determine the maximum eigenvalue $\lambda max$ as follows.

$$\boldsymbol{WV}=\begin{bmatrix} 1 & 1 & 1 & 6 & 1 \\ 1 & 1 & 0.2 & 6 & 1 \\ 1 & 5 & 1 & 7 & 1 \\ 0.167 & 0.167 & 0.143 & 1 & 0.167 \\ 1 & 1 & 1 & 6 & 1 \end{bmatrix}\begin{bmatrix} 0.2321 \\ 0.1682 \\ 0.3302 \\ 0.0375 \\ 0.2321 \end{bmatrix}=\begin{bmatrix} 1.1875 \\ 0.9233 \\ 1.8978 \\ 0.1901 \\ 1.1875 \end{bmatrix} \tag{10}$$

$$\lambda_{max}=\sum_{i=1}^{5}\frac{\left(WV\right)_i}{5V_i}=\frac{1.1875}{5\times0.2321}+\frac{0.9233}{5\times0.1682}+\frac{1.8978}{5\times0.3302}+\frac{0.1901}{5\times0.0375}+\frac{1.1875}{5\times0.2321}=5.3077 \tag{11}$$

The consistency index ($CI$) is

$$CI=\frac{\lambda_{max}-n}{n-1}=\frac{5.3077-5}{5-1}=0.0769\,. \tag{12}$$

The mean random consistency index ($RI$) obtained using Eq. (13) is determined by the AHP. When $n$ is 5, $RI$ is 1.12.

$$RI = 1.12 \tag{13}$$

$$CR = \frac{CI}{RI} = \frac{0.0769}{1.12} = 0.0687 \tag{14}$$

$$CR = 0.0687 < 0.1 \tag{15}$$

The AHP consistency ratio ($CR$) is shown in Eq. (14). When $CR$ is less than 0.1, as determined using Eq. (15), the consistency is acceptable in AHP.

By a similar method, we can calculate the weights of subindexes $\{I_{11}, I_{12}\}$ relative to the index $I_1$, the weights of subindexes $\{I_{21}, I_{22}\}$ relative to the index $I_2$, and so on. Table 2 shows all the relative and absolute weights of $L_2$ indexes.

For each expert response, we use AHP to calculate the weights of all indexes, as shown in Table 3. We can adjust the scores from the ten experts, and we calculate the average weights for the assessment process for each course.

Table 2
Indexes of assessment with weights.

| $L_0$ | $L_1$ | $L_2$ | Weight |
|---|---|---|---|
| Teaching quality | $I_1$ (0.2321): Content | $I_{11}$ (0.8571): Precise concepts, clear points and appropriate examples | 0.1989 |
| | | $I_{12}$ (0.1429): Combined theoretical and practical frontier development | 0.0332 |
| | $I_2$ (0.1682): Attitude | $I_{21}$ (0.8750): Well prepared, corrects exercises and answers questions | 0.1472 |
| | | $I_{22}$ (0.1250): Appropriate clothing, manner and punctuality | 0.0210 |
| | $I_3$ (0.3302): Skill | $I_{31}$ (0.8750): Modernized presentation and demonstration | 0.2889 |
| | | $I_{32}$ (0.1250): Precise and standard Chinese speaking and writing | 0.0413 |
| | $I_4$ (0.0375): Effect | $I_{41}$ (0.8571): Moderate difficulty and easy to understand | 0.0321 |
| | | $I_{42}$ (0.1429): High acceptance and interaction from students | 0.0054 |
| | $I_5$ (0.2321): Strategy | $I_{51}$ (0.5): Flexible, multiform and innovative | 0.1160 |
| | | $I_{52}$ (0.5): Inspires the students to think and learn | 0.1160 |

Table 3
Summary of assessment index weights.

| ExpertID | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $I_{11}$ | 0.1989 | 0.0935 | 0.1214 | 0.0569 | 0.178 | 0.0511 | 0.0569 | 0.0338 | 0.0919 | 0.1122 | 0.0995 |
| $I_{12}$ | 0.0332 | 0.0187 | 0.1214 | 0.0569 | 0.178 | 0.0511 | 0.0569 | 0.2365 | 0.0919 | 0.1122 | 0.0957 |
| $I_{21}$ | 0.1472 | 0.4447 | 0.1640 | 0.0444 | 0.1629 | 0.0511 | 0.2076 | 0.1714 | 0.1872 | 0.0489 | 0.1629 |
| $I_{22}$ | 0.021 | 0.0741 | 0.1640 | 0.0444 | 0.1629 | 0.0511 | 0.0415 | 0.0245 | 0.1872 | 0.0489 | 0.0819 |
| $I_{31}$ | 0.2889 | 0.0281 | 0.1108 | 0.1732 | 0.0227 | 0.0875 | 0.3801 | 0.1714 | 0.0919 | 0.0564 | 0.1411 |
| $I_{32}$ | 0.0413 | 0.007 | 0.0369 | 0.1732 | 0.1364 | 0.0113 | 0.0633 | 0.0245 | 0.0919 | 0.0081 | 0.0594 |
| $I_{41}$ | 0.0321 | 0.0662 | 0.0780 | 0.0477 | 0.0256 | 0.2695 | 0.0569 | 0.1714 | 0.0623 | 0.4584 | 0.1268 |
| $I_{42}$ | 0.0054 | 0.1987 | 0.0195 | 0.0477 | 0.0256 | 0.2695 | 0.0569 | 0.0245 | 0.0623 | 0.0573 | 0.0767 |
| $I_{51}$ | 0.116 | 0.0517 | 0.0920 | 0.1779 | 0.0135 | 0.0774 | 0.0685 | 0.0177 | 0.0666 | 0.0489 | 0.0730 |
| $I_{52}$ | 0.116 | 0.0172 | 0.0920 | 0.1779 | 0.0943 | 0.0774 | 0.0114 | 0.1242 | 0.0666 | 0.0489 | 0.0826 |

In the example, we use the TQA data from Sanda University. In the original data, there are more than 700 teachers. For each teacher, all students who have taken his/her courses fill out the questionnaire to evaluate the teaching quality of this teacher.

### 4.3. EDA–BP NN for evaluating grade

At the last stage of the TQA process, the EDA–BP NN is used to evaluate the grade of the teaching quality for each course. In this study, two variants of EDA are applied to the BP network: the PMBGA (classical EDA) and CGA. In addition, the traditional GA is still implemented for comparison. Table 4 shows the results of our work.

The CGA has the highest accuracy, precision, and F1 score. The training process of the EDA–BP NN is shown in Fig. 7. We can see that the CGA is the better algorithm with the lower convergence speed. On the other hand, the training loss of the EDA–BP NN is shown in Fig. 8. The CGA–BP is still the better algorithm. All the test set losses are smaller than the training set losses. This may indicate that the EDA applied to the BP network can effectively prevent overfitting. Figure 9 shows the accuracy of the training process. The accuracy is more meaningful than the loss function of the NN. We still find the CGA to be better than the others.

Table 4
Results of AHP–EDA–BP.

| Algorithm | Accuracy | Precision | Recall | F1 score |
|---|---|---|---|---|
| GA–BP | 0.9934 | 0.9826 | **0.9769** | 0.9700 |
| PMGA–BP | 0.9934 | 0.9600 | 0.9688 | 0.9194 |
| CGA–BP | **0.9956** | **0.9862** | 0.9695 | **0.9775** |

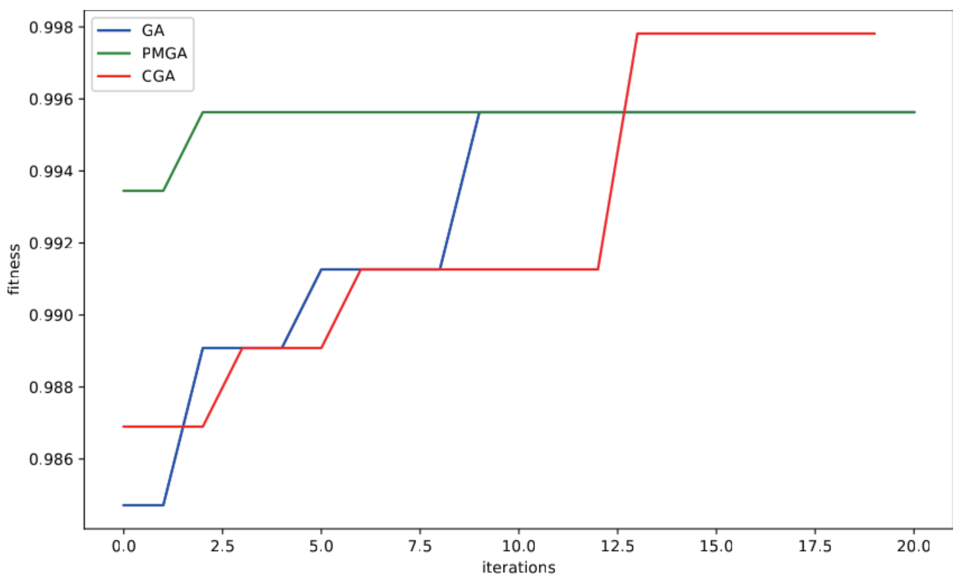



Fig. 7. (Color online) EDA–BP NN training fitness performance of three NNs: GA–BP, PMGA–BP, and CGA–BP.

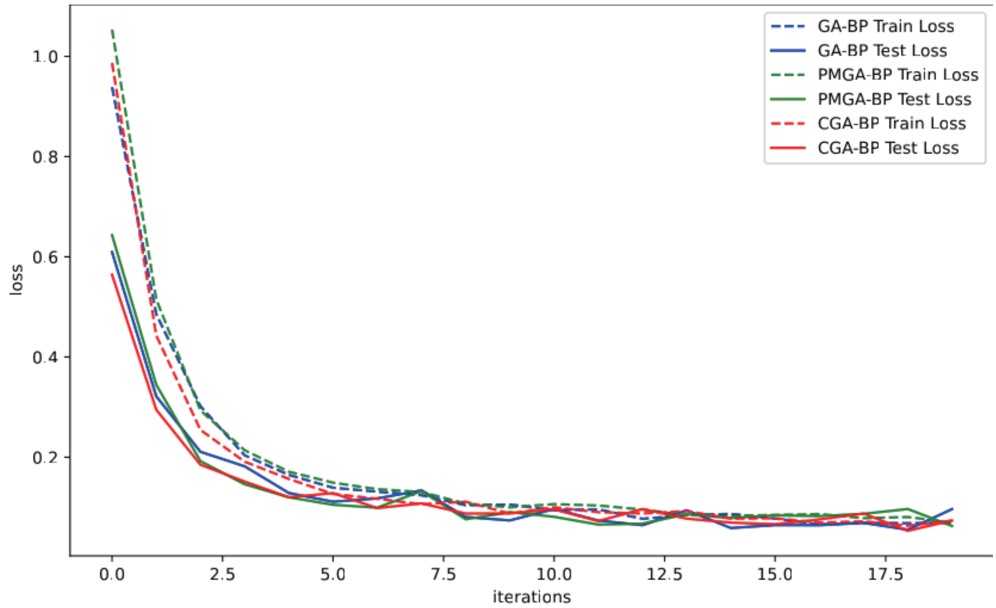



Fig. 8. (Color online) Training and testing set losses of the EDA–BP NNs GA–BP, PMGA–BP, and CGA–BP.

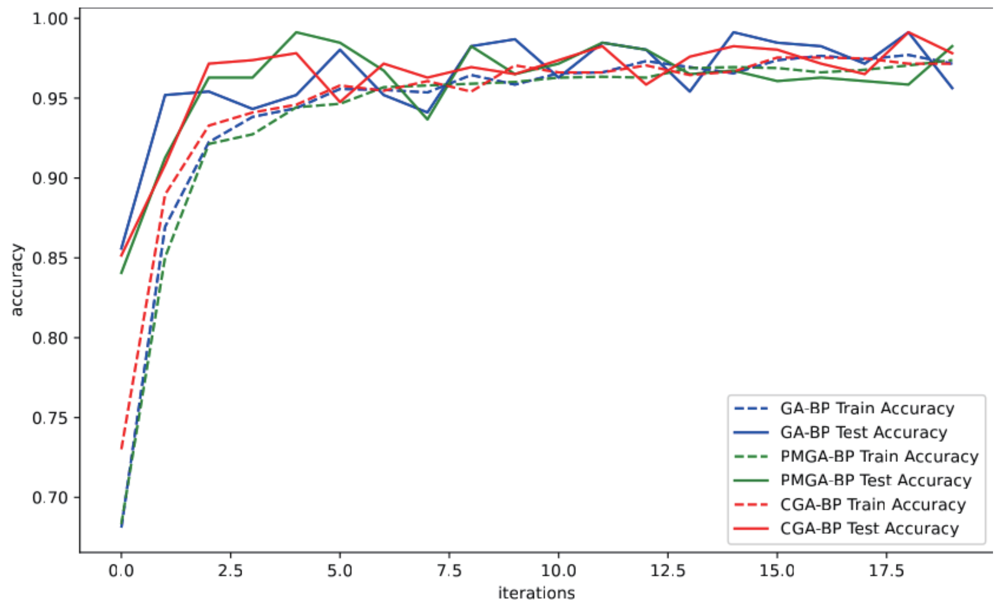



Fig. 9. (Color online) Accuracies of the training and testing sets of three EDA–BP NNs: GA–BP, PMGA–BP, and CGA–BP.

## 5. Implementation

Figure 10 shows the implementation of this prototype system. The implementation is based on the Django MVC framework and HTML5. Essentially, Python is used to write the

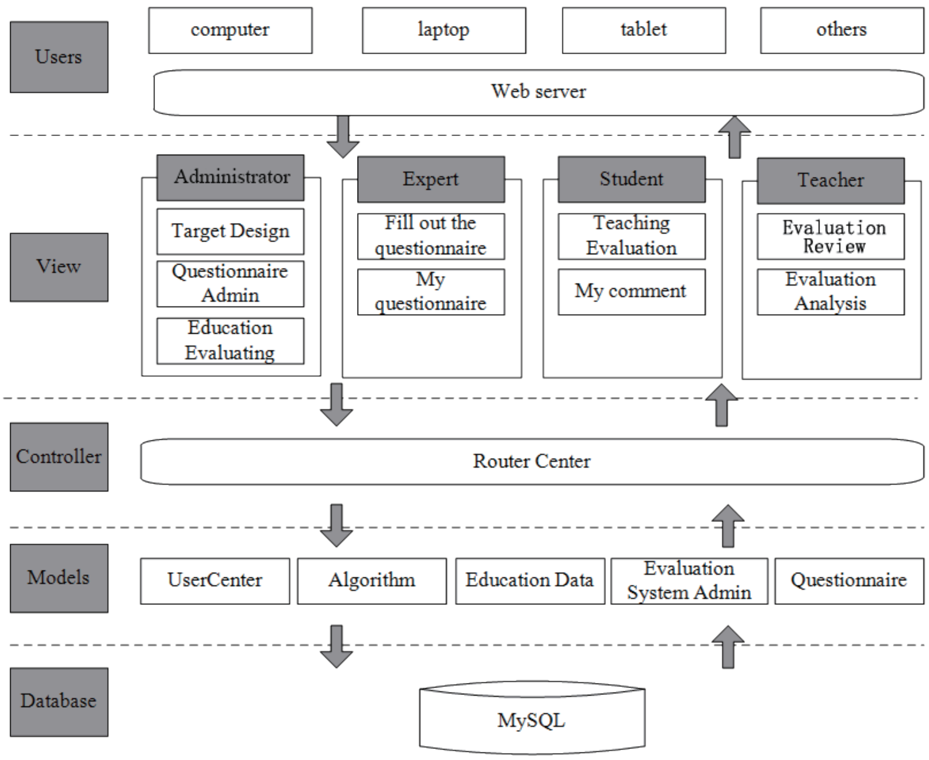



Fig. 10. The architecture of the TQA system follows the MVC architecture. The responsibilities of each layer and how the system allows user data to flow through the various levels of the MVC architecture and interact with the database are outlined.

system. The overall system is divided into three layers: the model, view, and control layers. The view layer provides the user interface and only transmits collected data to the control layer. The control layer manages and coordinates all system processes. After the control layer collects enough feedback, the control layer initiates the working flow process in the model layer. The model layer performs the analysis and database access operations. The results are returned to the view layer; then the administrators can acquire assessment results through the view layer. The system allows users to access the view layer through any commonly used HTML5 standard browser. The control layer is implemented on the powerful Django framework. At the same time, the Mysql database is used to manage all data in the TQA system.

The system has been applied to the TQA process at Sanda University, Shanghai, China. Figure 11 shows the questionnaire for experts to collect the relative weights of indexes. The index hierarchy and relative importance calculated via the expert questionnaires are shown in Fig. 12. In this system, the student can evaluate each course he/she takes, as shown in Fig. 13.


Fig. 11. (Color online) Expert's questionnaire. Experts can access the questionnaire by typing the obtained invitation code in the text box at the top of the page.


Fig. 12. (Color online) Administrator's view of the page showing indicator weights. The relative importance result was calculated using the responses from the teaching expert questionnaire.


Fig. 13. (Color online) Student's questionnaire in the TQA system. On this page, students can perform teaching evaluation operations.

## 6. Conclusions

Higher education institutions use the TQA system to measure teaching quality by collecting student feedback. How to use more intelligent methods to effectively analyze the feedback to indicate teaching effectiveness is the main research challenge at this time. In this study, we proposed and implemented a TQA system based on AHP–EDA–BP to illustrate the deeper meaning of student feedback. EDA can effectively transform the subjective scoring into quantitative values automatically. AHP–EDA–BP ensured the accuracy of the quantitative analysis. The experimental results show the essential applicability of the proposed method. In addition, the TQA system has been applied to the TQA process at Sanda University, Shanghai, China.

In the future, we will improve the availability and effectiveness of the proposed TQA system using the feedback from Sanda University. We will develop corresponding assessment criteria for different types of course. On the other hand, we will further develop AHP–EDA–BP to improve the accuracy of quantitative analysis. In this study, the experimental results and implementation of the TQA system showed the effectiveness and availability of sensor technology applied to higher education. However, our approach is still limited by the TQA system to only represent the short-term effect of teaching in higher education.

## Acknowledgments

This work was sponsored by the Shanghai Municipal Education Commission under contracts C2024057 (Educational Science Research Project) and Z90004.23.001, and the Key Courses Construction Project (Big data platform technology). This work was also supported by Sanda University under contract 2021ZD05.

## References

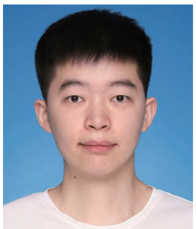
1 H. Bowen: Investment in Learning, H. Bowen, Ed. (Taylor & Francis, United Kingdom, 2018) 1st ed., pp. 349–353.
2 D. Vykydal, M. Folta, and J. Nenadál: Sustainability **12** (2020) 4769. https://doi.org/10.3390/su12114769
3 P. Feng: Front. Sport Res. **2** (2020) 2. https://doi.org/10.25236/fsr.2020.020413
4 S. Bianchini: Assess. Eval. Higher Educ. **39** (2013) 380. https://doi.org/10.1080/02602938.2013.842957
5 J. Dunlosky, K. A. Rawson, E. J. Marsh, M. J. Nathan, and D. T. Willingham: Psychol. Sci. Public Interest **14** (2013) 4. https://doi.org/10.1177/1529100612453266
6 M. S. Donovan and J. Bransford: How Students Learn, M. S. Donovan and J. Bransford, Eds. (National Academies Press, Washington, D.C., 2005) 1st ed.
7 T. Ryan: Higher Learn. Res. Commun. **5** (2015). https://doi.org/10.18870/hlrc.v5i4.257
8 Z. J. Liu and S. P. Tseng: Proc. 2020 8th Int. Conf. Orange Technology (IEEE, 2020) 1–3.
9 X. Guo, M. Gao, M. Zhang, Y. Chen, and S.-P. Tseng: Proc. 2020 8th Int. Conf. Orange Technology (IEEE, 2020) 1–3.
10 R. K. F. Ip, S. I. F. Iong, M. X. Y. Wu, and S. S. Y. Wang: Proc. 2017 IEEE 6th Int. Conf. Teaching, Assessment, and Learning for Engineering (IEEE, 2017) 444–448.
11 D. Feistauer and T. Richter: Assess. Eval. Higher Educ. **42** (2016) 1263. https://doi.org/10.1080/02602938.2016.1261083
12 M. Li and Y. Su: Int. J. Emerging Technol. Learn. (iJET) **15** (2020) 147. https://doi.org/10.3991/ijet.v15i16.15937

13 A. R. Mishra, D. Jain, and D. S. Hooda: Proc. 2nd Int. Conf. Computer and Communication Technologies (IC3T, 2015) 387–339.
14 Y. He, Z. Meng, H. Xu, and Y. Zou: Physica A **556** (2020) 124845. https://doi.org/10.1016/j.physa.2020.124845
15 C. Blum and A. Roli: ACM Comput. Surv. **35** (2003) 268. https://doi.org/10.1145/937503.937505
16 S. Wang, C. Yu, D. Shi, and X. Sun: Math. Probl. Eng. **2018** (2018) 1. https://doi.org/10.1155/2018/6483145
17 J. Douglas and A. Douglas: Qual. Higher Education **12** (2006) 3. https://doi.org/10.1080/13538320600685024
18 N. Shi and Y. Hu: Int. J. Technol. Inclusive Educ. **3** (2014) 248. https://doi.org/10.20533/ijtie.2047.0533.2014.0032
19 M. Goos and A. Salomons: Res. Higher Educ. **58** (2016) 341. https://doi.org/10.1007/s11162-016-9429-8
20 W. E. Becker and M. Watts: Am. Econ. Rev. **89** (1999) 344. https://doi.org/10.1257/aer.89.2.344
21 S. O. Michael: Int. J. Educ. Manage. **11** (1997) 117. https://doi.org/10.1108/09513549710164014
22 P. A. Pasque, L. A. Hendricks, and N. A. Bowman: Proc. National Forum on Higher Education for the Public Good (ERIC, 2006).
23 L. P. Symaco and M. Y. Tee: Int. J. Educ. Dev. **66** (2019) 184. https://doi.org/10.1016/j.ijedudev.2018.10.001
24 D. A. McManus: J. Geosci. Educ. **49** (2001) 423. https://doi.org/10.5408/1089-9995-49.5.423
25 L. Lomas and G. Nicholls: Qual. Higher Educ. **11** (2005) 137. https://doi.org/10.1080/13538320500175118
26 L. R. Goldberg, D. F. Parham, K. L. Coufal, M. Maeda, R. R. Scudder, and P. R. Sechtem: J. College Teach. Learn. (TLC) **7** (2010) 71. https://doi.org/10.19030/tlc.v7i2.91
27 S. Samson and D. E. McCrea: Ref. Serv. Rev. **36** (2008) 61. https://doi.org/10.1108/00907320810852032
28 H. Hu and J. Zheng: Int. J. Emerging Technologies in Learning (iJET) **11** (2016) 16. https://doi.org/10.3991/ijet.v11i08.6040
29 T. L. Saaty: Eur. J. Oper. Res. **48** (1990) 9. https://doi.org/10.1016/0377-2217(90)90057-i
30 T. L. Saaty: Rev. R. Acad. Cienc. Exactas, Fis. Nat. Ser. A Mat. **102** (2008) 251. https://doi.org/10.1007/bf03191825
31 T. L. Saaty: J. Adv. Transp. **29** (1995) 81. https://doi.org/10.1002/atr.5670290109
32 M. Kurttila, M. Pesonen, J. Kangas, and M. Kajanus: For. Policy Econ. **1** (2000) 41. https://doi.org/10.1016/s1389-9341(99)00004-0
33 O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad: Heliyon **4** (2018) e00938. https://doi.org/10.1016/j.heliyon.2018.e00938
34 S. Shanmuganathan: Artificial Neural Network Modelling: An Introduction, S. Shanmuganathan and S. Samarasinghe, Eds. (Springer, 2016) 1st., pp. 1–14.
35 W. Samek, G. Montavon, S. Lapuschkin, C. J. Anders, and K.-R. Muller: Proc. the IEEE 2021. (IEEE, 2021) 247–278.
36 D. E. Rumelhart, R. Durbin, R. Golden, and Y. Chauvin: Backpropagation (Psychology Press, 2013) pp. 1–34.
37 D. E. Rumelhart, G. E. Hinton, and R. J. Williams: Nature **323** (1986) 533. https://doi.org/10.1038/323533a0
38 S. Amari: Neurocomputing **5** (1993) 185. https://doi.org/10.1016/0925-2312(93)90006-o
39 K. Hussain, M. N. Mohd Salleh, S. Cheng, and Y. Shi: Artif. Intell. Rev. **52** (2018) 2191. https://doi.org/10.1007/s10462-017-9605-z
40 S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi: Science **220** (1983) 671. https://doi.org/10.1126/science.220.4598.671
41 J. H. Holland: Sci. Am. **267** (1992) 66. https://www.jstor.org/stable/24939139
42 M. R. Bonyadi and Z. Michalewicz: Evol. Comput. **25** (2017) 1. https://doi.org/10.1162/evco_r_00180
43 M. Hauschild and M. Pelikan: Swarm Evol. Comput. **1** (2011) 111. https://doi.org/10.1016/j.swevo.2011.08.003
44 M. Pelikan, M. W. Hauschild, and F. G. Lobo: Springer Handbook of Computational Intelligence (2015) p. 899. https://doi.org/10.1007/978-3-662-43505-2_45
45 M. Pelikan and M. Pelikan: Hierarchical Bayesian Optimization Algorithm. (Springer Berlin Heidelberg, 2005) 1st., pp. 105–129.
46 G. R. Harik, F. G. Lobo, and D. E. Goldberg: IEEE Trans. Evol. Comput. **3** (1999) 287. https://doi.org/10.1109/4235.797971.

## About the Authors

**Xin Guo** received her B.E. degree from the Department of Computer Science and Technology, Henan Institute of Science and Technology, China, in 2005 and her M.Ed. degree from the Department of Educational Technology, Shanghai Normal University, China, in 2008. From 2008 to 2024, she was an associate professor at Sanda University, China. Since 2024, she has been an associate professor at Shanghai Pudong Vocational and Technical College. Her research interests are in data mining and analysis, machine learning, and educational technology. (smilegx110@hotmail.com)

**Ying Chen** received her B.S. degree from Fudan University, China, in 1990, her M.M. degree from Shanghai Second Medical University in 1998, and her Ph.D. degree from Shanghai Jiao Tong University in 2006. From 2013 to 2014, she was an associate professor at Shanghai Jiao Tong University, China. Since 2014, she has been a professor at Sanda University. Her research interests are in data engineering, intelligent technologies and applications, and biomedical engineering. (ychen@sandau.edu.cn)



**Zilong Yin** received his Master's degree in electronic information from Shanghai University of Engineering Science in 2025. He is currently pursuing his Ph.D. in software engineering at East China Normal University in Shanghai. His research interests include privacy computing, federated learning, and large language models. (zilong_yin@163.com)

**Ruoying Wang** received her B.S. degree from Sanda University, China, in 2023 and her M.S. degree from King's College London, UK, in 2024. Her research interests are in data science and bioengineering. (ruoyingwang2000@outlook.com)

**Dazhi Li** received his B.S. degree from Northwest Normal University in 2002, his M.S. degree from Shanghai Normal University in 2008, and his Ph.D. degree from Shanghai Jiao Tong University in 2018. Since 2002, he has been serving as a lecturer in the Department of Computer Science at Shanghai Normal University. His research interests focus on the Internet of Things (IoT), cloud computing, and the management of educational informatization. (lijunzhi@shnu.edu.cn )

**Shih-Pang Tseng** received his B.S. and M.S. degrees from the Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan, and his Ph.D. degree from the Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan. At present, he is a professor at the School of Software and Big Data, Changzhou College of Information Technology, Changzhou, China. His current research interests include artificial intelligence, learning technology, the Internet of Things, and robotics. (tsengshihpang@czcit.edu.cn)