A new method for education quality evaluation based on belief rule base with power set and evidential reasoning

MinJie Liang
Harbin Normal University
Guohui Zhou (zhouguohui@hrbnu.edu.cn)
Harbin Normal University
Wei He
Harbin Normal University
Haobing Chen
Harbin Normal University
Jidong Qian
Harbin Normal University

Research Article

Keywords: education quality evaluation, evidential reasoning, belief rule base, power set, parameters optimization

Posted Date: December 20th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3757735/v1

Additional Declarations: No competing interests reported.

A new method for education quality evaluation based on belief rule base with power set and evidential reasoning

Abstract

The evaluation of education quality is an important part of the construction of the education system, and it is a direct path to improving education quality. Education quality is a complex evaluation system that takes into account a number of dimensions, such as objectives, attitudes and outcomes. In teaching practice, these dimensions are often good or bad, making it difficult to assess the results of a comprehensive evaluation, leading to the problem of ignorance. Therefore, a belief rule base with power set (PBRB) is used to construct the model for education quality evaluation. The model extends the set of evaluation results into a power set that expresses a variety of evaluation ratings that are difficult to characterize, and uses the BRB to assign belief degrees to each rating in the power set. In addition, because the evaluation of education quality requires a great number of indicators, the BRB faces the rule combination explosion problem. To solve the problem, the transformation matrix is constructed to unify the evaluation indicators, which are then fused using the evidential reasoning (ER) algorithm to reduce the number of PBRB attributes. Finally, the parameters are optimized with the projected covariance matrix adaptive evolution strategy (P-CMA-ES) to improve the model accuracy. According to the experimental results, compared with other methods, the education quality evaluation method based on the EPBRB proposed in this paper can give a clear evaluation grade and has better accuracy and stability.

Keywords: education quality evaluation; evidential reasoning; belief rule base; power set; parameters optimization

1. Introduction

Education quality evaluation is the process of measuring the teaching process using scientific methods and assigning a value to the results. It focuses on specific teaching objectives and is guided by certain educational values [1]. With the increasing scale of education, the evaluation of education quality has also received more attention and has become an important part of the construction of a modern education system [2]. However, education quality itself is an abstract concept and is not easy to define and evaluate. In addition, the evaluation results need to be objective and credible [3]. Therefore, an evaluation method with clear definitions of various parameters and strong reliability and interpretability is needed.

In recent years, researchers in the education field have proposed many evaluation methods. Fan et al. established an evaluation system for the quality of innovative education on the basis of the CIPP model, and then applied the AHP and FCE to establish a fuzzy evaluation model. The method used AHP to determine the indicator weights and FCE to evaluate the effect of education, which realized the full integration of quantitative and qualitative information [4]. In the context of online education, Li and Su found that the evaluation indicators at different levels have different importance to the evaluation results. They used the entropy weighting method to process the indicator weights, and the gray cluster analysis method was used to analyze the education quality ratings of different evaluation objects [5]. Lv proposed an education quality evaluation method based on data mining, which designed the mining process and specific classification in the analysis of association rules. Through experiments, this method can derive the relationship between teacher competence, teaching methods and teaching effectiveness [6]. Wang combined data mining and IoT telecommunications to enable the collection, organization, processing and analysis of educational data [7]. Liu et al. proposed a method for evaluating the quality of undergraduate education, which was based on the back propagation (BP) neural network and stress test. They selected 19 indicators to assess education quality and then designed a BP neural network to construct the model. Finally, the indicators were examined for sensitivity using stress test [8]. Niu analyzed the defects and reasons of the fuzzy comprehensive evaluation method in the evaluation of preschool education quality and proposed an evaluation method combined with feed forward neural network (FNN). The method not only improved the expression ability of fuzzy information but also gave full play to the learning ability of neural networks [9]. Fang designed a portable wearable device to acquire students' ECG signals, analyzed the data and extracted features from the QRS wave using a one-dimensional convolutional neural network ((1D-CNN) model, and finally trained an SVM classifier with three levels of intensity. Taking physical condition and exercise intensity, etc. As a judgment factor, it is used to improve the quality of physical education [10]. Yuan proposed a fuzzy evaluation model of education quality based on the Markov chain. The authors used evaluations from students in a major course as experimental data and designed a corresponding hybrid education quality evaluation model [11]. Fouskakis proposed a Bayesian beta regression model with a Dirichlet prior, which enables the quantification of the degree of student preference in courses. The model has a priori information fusion capability to continuously monitor and evaluate education quality [12].

The methods briefly described above can be broadly categorized into three groups. Although all of these methods have been successfully applied to education quality evaluation, all of them have some limitations. Firstly, regarding the model analysis method. The weights of the relevant indicators depend excessively on subjective factors to determine, and the learning ability of the model is poor, resulting in evaluation results that are not sufficiently accurate and objective [13]. For the secondary approaches, the data-driven model needs enough data for training to obtain high accuracy, so it is not applicable in small-sample conditions. Due to the many uncertainties associated with data collection in the education field, it is difficult to obtain sufficient valid data, which makes these approaches not applicable to some practical situations [14]. More critically, most AI algorithms are not interpretable at all, causing the evaluation results to be unconvincing [15]. Multiple factors are considered in the hybrid information models, and the computational process is complicated when the volume of data is large. To overcome the problems in the existing methods, BRB is introduced to construct the education quality evaluation model in this paper.

The BRB method was proposed by Yang et al. in 2006 [16], which is based on the traditional if-than rule and embeds the belief degree into the result of the rule, so it can express various types of uncertainty information [17]. The BRB can accept both quantitative and qualitative information to construct the initial rule, and the initial parameter can be determined by expert knowledge [18]. Moreover, BRB has the ability to be optimized by training to improve accuracy. It is essentially a white-box model with a completely transparent inference process and strong interpretability of the results [19]. BRBs have been implemented in the fields of fault detection [20], risk assessment [21], industrial alarms [22] and disease diagnosis [23].

However, the system of education quality evaluation is complex with numerous indicators, which will lead to too many belief rules in constructing the BRB, so the ER algorithm is used to simplify the rules. In this paper, an evaluation system with two levels is constructed. Firstly, the transformation matrices are used to construct the mapping relationship, which transforms a set of secondary indicators into the belief distribution under the corresponding first-level indicators. Then, these probability vectors in the same space are synthesized with the ER algorithm, and the fusion result is put into the BRB, which can effectively reduce the number of antecedent attributes [24]. Additionally, some similar indicator data may lead to ambiguities between different evaluation results and cannot clearly generate a good or bad result. Such problems include local ignorance and global ignorance. The original BRB cannot properly assign belief degrees when dealing with this type of problem, so the power set of evaluation results is used as an identification framework, allowing the model to express more complex information [25].

The innovations in this paper are as follows:

(1) A new EPBRB model is designed. The matrix is used in this model to achieve the data transformation, and the ER algorithm is used to fuse the indicators. With this approach, the number of antecedent attributes of the BRB can be reduced reasonably to solve the problem of rule combination explosion. For the posteriors of the BRB, the identification framework is extended to a power set so that the BRB can express uncertain results to solve the ignorance problem.

(2) The EPBRB is applied to education quality evaluation for the first time. In this paper, a hierarchical education quality evaluation system is constructed, and the relationships among evaluation indicators are sorted. On this basis, the mapping between indicators of different levels is constructed using the transformation matrix to complete the information transformation. Then, the second-level indicators are fused using the ER algorithm to obtain the input of the PBRB. Finally, the belief degrees are assigned to the power set of all evaluation results to derive the final education quality rating.

The remainder of the paper is structured as follows. In Section 2, the problems in education quality evaluation are analyzed and solved. In Section 3, a model of education quality evaluation based on EPBRB is constructed. In Section 4, a case study is designed to validate the model performance. In Section 5, the research work of this paper is summarized and the aspects that can be improved are analyzed.

2. Problem formulation

In Section 2.1, problems in education quality evaluation are analyzed. In Section 2.2, an education quality evaluation model based on EPBRB is proposed.

2.1 Problem formulation of education quality evaluation

Education quality evaluation data are mostly obtained by questionnaires in the form of scales, covering multiple dimensions and containing a large number of indicators, which requires evaluation methods to have the ability to nonlinearly model complex systems. The accuracy and interpretability of the model are more meaningful due to the sensitivity of the evaluation results in the education field. The main problems of the education quality evaluation model based on EPBRB are summarized as follows:

Problem 1. Processing of evaluation indicator data. Education quality needs to be evaluated from many dimensions and involves many evaluation indicators. Modeling directly with data from these indicators would make the BRB evaluation model too complex [26]. The education quality evaluation system designed in this paper has two levels, and there is a certain relationship between the indicators of different levels, so the transformation matrices are used to construct the mapping relationship between them. First, the indicators in each level are set to the corresponding reference grades, and then the transformation matrices are used to construct the mapping relationship from the secondary indicators to the first-level indicators. The mapping relationship is constructed by

Equation (1):

$$(R_{1,i}, R_{2,i}, ..., R_{p,i}) = T_i(G_{1,i}, G_{2,i}, ..., G_{P_i,i}) \tag{1}$$

where $R_{1,i}, R_{2,i}, \ldots, R_{P,i}$ denotes the P reference grades of the ith first-level indicator, below which there are several secondary indicators, and $G_{1,i}, G_{2,i}, \ldots, G_{p_i,i}$ denotes the P_i reference grades of the secondary indicators. $T_i(\cdot)$ denotes the process of mapping using the transformation matrix.

Problem 2. The identification framework is extended to a power set so that the model can express both local ignorance and global ignorance. The identification framework of the original BRB is defined as $\Omega = \{D_1, D_2, \ldots, D_N\}$, where D_i denotes the ith evaluation result. The original BRB model cannot give the evaluation result explicitly when the inputs of the indicators are similar. However, after the extension of the identification framework to power sets, all possible results can be expressed, and the evaluation results are more precise. In education quality evaluation, local ignorance means that the evaluation result is not sure which one of the partial results it is, and global ignorance means that the evaluation result may be any one of the N results. The power set identification framework is expressed as Equation (3):

$$2^{\Omega} = \{\varnothing, D_1, ..., D_N, \{D_1, D_2\}, \{D_1, D_3\}, ..., \{D_1, ..., D_{N-1}\}, \Omega\}, S_n \subseteq 2^{\Omega} \tag{2}$$

where $\emptyset$ is the empty set. $\{D_i, D_j\}$ denotes that the evaluation result may be D_i or D_j and is used to express local ignorance. Ω is the full set of all results and is used to express global ignorance. S_n denotes the nth evaluation rating, such as $S_2 = D_1$ and $S_{2^N} = \Omega$.

Problem 3. Construction and optimization process of the EPBRB model. The input data are transformed into belief distributions by the processing of the transformation matrix. These belief distributions are used as evidence and fused using the ER algorithm. The obtained results are used as inputs to the PBRB to reduce the number of attributes [27]. Then, the PBRB model is constructed, and inference is performed to obtain the evaluation results. The EPBRB model can be constructed by combining expert knowledge, but incomplete expert knowledge is not accurate enough to affect the model performance. The optimization process is constructed in this paper based on P-CMA-ES, and the parameters in the ER and PBRB are optimized simultaneously. The construction of the model is represented by Equations (3) and (4), and the parameter optimization is represented by Equation (5):

$$y_i = ER(t(x_1), t(x_2), ..., t(x_J), V) \tag{3}$$

$$u(S(y)) = PBRB(y_1, y_2, ..., y_M, Q) \tag{4}$$

$$\Phi = O(EPBRB(\cdot)) \tag{5}$$

where y_i denotes the result of the ER algorithm. $ER(\cdot)$ denotes the process of fusing J pieces of evidence by the ER algorithm. $t(x_1), t(x_2) \ldots, t(x_J)$ denotes a set of belief distributions obtained from the input indicators. V is the set of parameters of the ER algorithm. $u(\cdot)$ denotes the result of the education quality evaluation. $S(\cdot)$ is the grade of education quality. $PBRB(\cdot)$ denotes the construction of the PRBB model. $y_1, y_2, \ldots, y_M$ denotes the fusion result as an attribute input. Q is the set of parameters of the PBRB model. Φ is the optimized set of optimal parameters. $O(\cdot)$ denotes the process of global optimization of parameters.

2.2 Construction of the education quality evaluation model

Regarding the above problems, in this section, a model for education quality evaluation based on EPBRB is constructed. The model fuses complex indicators with the ER algorithm and uses belief rules as the inference basis, and the inference results are the belief degrees assigned to the power set identification framework. The rules of the EPBRB model are shown in Equation (6):

$$\begin{aligned} R_k &: IF (y_1 \ is \ A_1^k) \wedge (y_2 \ is \ A_2^k) \wedge \ldots \wedge (y_M \ is \ A_M^k) \\ & THEN \ \{(S_1, \beta_{1,k}), (S_2, \beta_{2,k}), \ldots, (\Omega, \beta_{2^N,k})\}, \left(\sum_{n=1}^{2^N} \beta_{n,k} = 1 \right) \\ & WITH \text{ rule weight } \theta_k \\ & AND \text{ attribute weight } \delta_1, \delta_2, \ldots, \delta_M \end{aligned} \tag{6}$$

where $R_k (k = 1,2, \ldots, L)$ is the kth rule in EPBRB and L is the total number of rules. $y_1, y_2, \ldots, y_M$ are M fused evaluation indicators. $A_1^k, A_2^k, \ldots, A_M^k$ denote the reference values of the model attributes. When the input value of attribute y_i is A_i^K, the rule derives a belief degree $\beta_{n,k}$ corresponding to each outcome of Equation (3). θ_k is the rule weight of R_k. δ_i is the attribute weight.

The components of the EPBRB model are as follows. First, the secondary indicators are mapped as vectors in the same space using the transformation matrix, and then the indicators are fused by the ER algorithm. Second, the identification framework is extended as a power set to construct the EPBRB model. Third, the parameters in the EPBRB are optimized by the optimization algorithm to improve the accuracy of the model results. The structure of the model is shown in Figure 1.

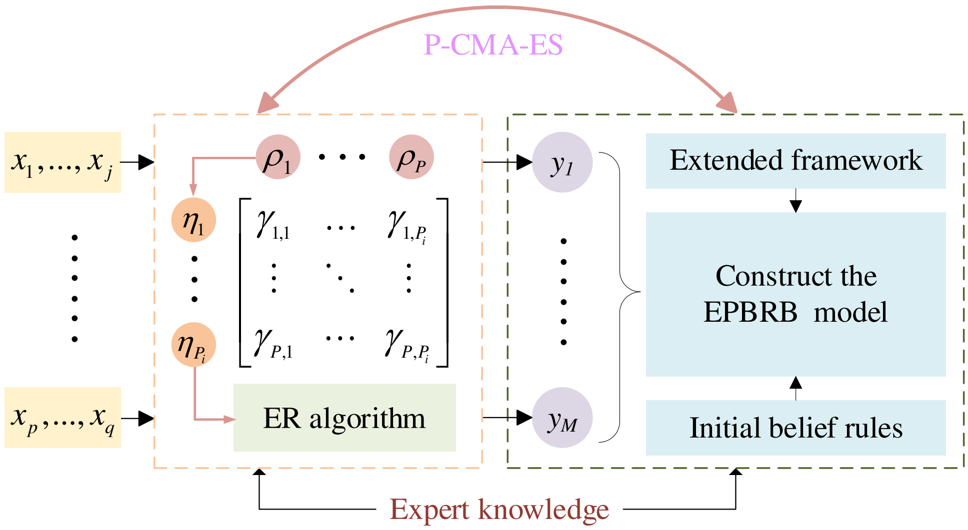

Figure 1. Structure of the EPBRB model

3. Education quality evaluation model based on EPBRB

In this section, the modeling process of the education quality evaluation model based on EPBRB is specifically described. In Section 3.1, the method of mapping between indicators at different levels is realized with a transformation matrix. The inference process of the EPBRB model is shown in Section 3.2. In Section 3.3, the optimization process based on P-CMA-ES is constructed. In Section 3.4, the structure of the education quality evaluation model based on EPBRB is demonstrated.

3.1 Indicator transformation based on matrix

The evaluation system in this paper has two levels: the first-level indicators are the dimensions for evaluating education quality, and the secondary indicators are more detailed divisions. In the example of education quality evaluation, the data for the secondary indicators are obtained from the scale, which is set up with a specific level of data, different from the standard of the first-level indicators. Therefore, the reference grades of indicators at different levels are not consistent. When data from different indicators are used as inputs to the evaluation model, they can be transformed into a match of the same attribute reference value by the reference value set for each indicator. This process is constructed based on the transformation matrix:

Suppose the ith first-level indicator Y_i is set with P reference grades, denoted as $\{R_{k,i}|k=1,2,\ldots,P\}$. There are N secondary indicators below it, set with P_i reference grades, denoted as $\{G_{k,i}|k=1,2,\ldots,P_i\}$. The transformation process is shown in Figure 2.

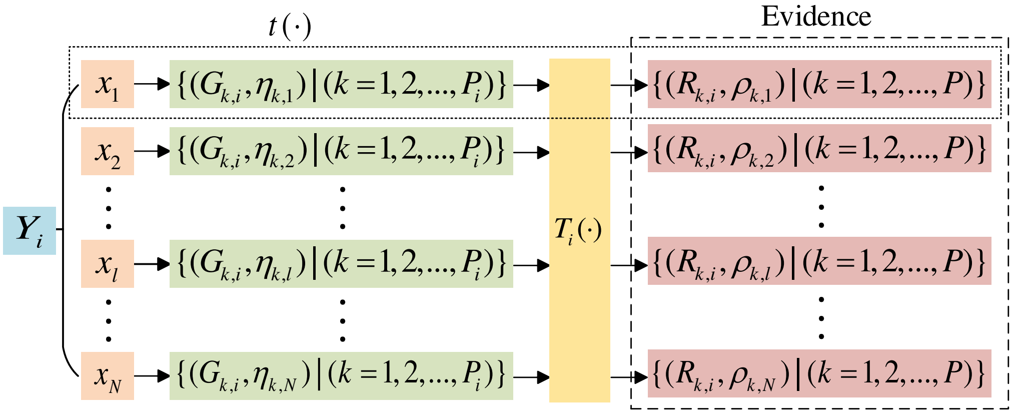

Figure 2. Transformation process between indicators at different levels

The $t(\cdot)$ in the above figure is the process of transforming an input x_l in Equation (2). First, the input is transformed into a belief distribution of $\{G_{k,i}\}$ using a data transformation method. Then, the transformation matrix is constructed based on expert knowledge. Finally, the belief distribution of the secondary indicator is mapped to the belief distribution of $\{R_{k,i}\}$, which is a piece of evidence in the ER algorithm. If the probabilities in the belief distribution are considered vectors, then the process is essentially a mapping of a vector to another space through a matrix, which can be represented by Equation (7):

$$(\eta_{1,l},\eta_{2,l},\ldots,\eta_{P_i,l})^T \xrightarrow{T_i(\cdot)} (\rho_{1,l},\rho_{2,l},\ldots,\rho_{P,l})^T,(\sum_{k=1}^{P_i}\eta_{k,l}=1,0\le\eta_{k,l},\rho_{k,l}\le 1) \tag{7}$$

where $\eta_{k,l}$ is the belief degree that the input value belongs to the reference grade $G_{k,i}$. $\rho_{k,l}$ is the belief degree that the value of the first-level indicator belongs to the reference grade $R_{k,i}$.

Each input is converted into a belief distribution about the reference grades based on the information transformation method, as shown in Equation (8).

$$\begin{cases} \eta_{k,l}=\dfrac{G_{k+1,l}-x_l}{G_{k+1,l}-G_{k,l}} & G_{k,l}\le x_l\le G_{k+1,l} \\ \eta_{k+1,l}=1-\eta_{k,l} & G_{k,l}\le x_l\le G_{k+1,l} \\ \eta_{q,l}=0 & q=1,\ldots,P_i,q\ne k,k+1 \end{cases} \tag{8}$$

where $G_{k+1,l}$ and $G_{k,l}$ are the two closest reference values to x_l.

$T_i(\cdot)$ in Figure 2 is the process shown in Equation (1). The transformation matrix is constructed as shown in Equation (9):

$$T_i = \begin{array}{c} \\ R_{1,i} \\ R_{2,i} \\ \vdots \\ R_{P,i} \end{array} \begin{array}{c} \begin{array}{cccc} G_{1,i} & G_{2,i} & \cdots & G_{P_i,i} \end{array} \\ \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & \cdots & \gamma_{1,P_i} \\ \gamma_{2,1} & \gamma_{2,2} & \cdots & \gamma_{2,P_i} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{P,1} & \gamma_{P,2} & \cdots & \gamma_{P,P_i} \end{bmatrix} \end{array} (\sum_{k=1}^{P} \gamma_{k,n} = 1, n = 1,2,\ldots,P_i) \tag{9}$$

where T_i is the transformation matrix mapping to the reference grades of the ith first-level indicator. Each of its columns is a set of belief distributions, where each element denotes the belief degree that $G_{k,i}$ belongs to $R_{k,i}$, denoted as $G_{k,i} = \{(R_{1,i}, \gamma_{1,k}), (R_{2,i}, \gamma_{2,k}), \ldots, (R_{P,i}, \gamma_{P,k})\}$. γ is given by expert knowledge, and the sum of the elements in each column is equal to 1.

Finally, $\rho_{k,l}$ is obtained from the transformation matrix and the belief distribution of the reference grades of the secondary indicators. The calculation is shown in Equation (10):

$$\xi_l = T_i \times \vartheta_l, \rho_{R,l} = 1 - \sum_{n=1}^{P} \xi_{n,l} \tag{10}$$

where $\xi_l = (\rho_{1,l}, \rho_{2,l}, \ldots, \rho_{P,l})^T$ is the belief distribution about the reference grades of the first-level indicators obtained by matrix transformation. $\vartheta_l = (\eta_{1,l}, \eta_{2,l}, \ldots, \eta_{P_i,l})^T$ is the belief distribution of the input x_l according to Equation (8). $\rho_{R,l}$ is the probability unassigned to the set $\{R_{k,i}\}$, which denotes global ignorance.

3.2 Inference process of evaluation model based on EPBRB

The input data are transformed to give different belief degrees for the same evaluation dimension, and the ER algorithm can fuse these results to produce a comprehensive evaluation. The fusion of each set of secondary indicators will result in data for the first-level indicators. The attributes of the PBRB model are reduced after fusion, while the number of rules is greatly reduced.

The inference of EPBRB starts with indicator fusion. The model receives a set of data from secondary indicators, first transforms them into the form of a belief distribution, and then transforms them into a different description of the first-level indicators based on the transformation matrix. As a result of this process, each input is transformed into a piece of belief distribution about a first-level indicator, which is regarded as evidence. Finally, the initial evidence weight is set $q_i (0 \leq q_i \leq 1)$. The fusion process of the ER algorithm is as follows:

$$\begin{gathered} \varpi_n = \frac{\upsilon \left[\prod_{k=1}^{N} (q_k \rho_{n,k} + 1 - q_k \sum_{j=1}^{P} \rho_{j,k}) - \prod_{k=1}^{N} (1 - q_k \sum_{j=1}^{P} \rho_{j,k}) \right]}{1 - \upsilon \left[\prod_{k=1}^{J} (1 - q_k) \right]} \\ \upsilon = \frac{1}{\sum_{n=1}^{P} \prod_{k=1}^{N} (q_k \rho_{n,k} + 1 - q_k \sum_{j=1}^{P} \rho_{j,k}) - (N-1) \prod_{k=1}^{N} (1 - q_k \sum_{j=1}^{P} \rho_{j,k})} \end{gathered} \tag{11}$$

where ϖ_n is the belief degree of the nth reference level R_n of the first-level indicator that satisfies $0 \leq \varpi_n \leq 1$ and $\sum_{n=1}^{P} \varpi_n = 1$. $\rho_{j,k}$ is obtained from Equation (10).

Suppose the utility value of the indicator reference grade R_n is $u(R_n)$; then, the expected utility of the first-level indicator is calculated as shown in Equation (12):

$$y_i = \sum_{n=1}^{P} u(R_n)\varphi_n \tag{12}$$

where y_i is the input of the PBRB obtained by the ER algorithm.

The attributes of the evaluation model have been rationally reduced after fusing the indicators by the ER algorithm. The next step in the inference process is to calculate the matching degree of the inputs in the PBRB for each reference value. It is shown by the following equation:

$$a_i^k = \begin{cases} \dfrac{A_i^{l+1} - y_i}{A_i^{l+1} - A_i^l} & k = l(A_i^l \le y_i \le A_i^{l+1}) \\ \dfrac{y_i - A_i^l}{A_i^{l+1} - A_i^l} & k = l+1 \\ 0 & k = 1,2,\ldots,L(k \ne l, l+1) \end{cases} \tag{13}$$

where a_i^k is the matching degree of the indicator data to the reference value A_i. y_i is the value from Equation (12) as an input to the PBRB. A_i^l and A_i^{l+1} are the two neighboring reference values of the ith antecedent attribute.

After obtaining the matching degree, the rules whose matching degree is not all 0 are activated. Next, the activation weight is calculated, which indicates how much the rule is activated. It is calculated by Equation (14):

$$\omega_k = \frac{\theta_k \prod_{i=1}^{M} (a_i^k)^{\overline{\delta_i}}}{\sum_{l=1}^{L} \theta_l \prod_{i=1}^{M} (a_i^l)^{\overline{\delta_i}}}, \overline{\delta_i} = \frac{\delta_i}{\substack{\max\{\delta_i\} \\ i = 1,2,\ldots,M}} \tag{14}$$

where ω_k denotes the activation weight of the kth rule. θ_k denotes the weight of the kth rule. δ_i denotes the weight of the ith antecedent attribute, and $\overline{\delta_i}$ denotes the relative weight of the attribute. M is the number of antecedent attributes.

The activated rules are then synthesized using the ER algorithm to produce the results of the model. The ER algorithm is shown in Equation (15):

$$\begin{gathered} \beta_n = \frac{\mu\left[\prod_{k=1}^{L}(\omega_k\beta_{n,k} + 1 - \omega_k\sum_{j=1}^{2^N}\beta_{j,k}) - \prod_{k=1}^{L}(1 - \omega_k\sum_{j=1}^{2^N}\beta_{j,k})\right]}{1 - \mu\left[\prod_{k=1}^{L}(1-\omega_k)\right]} \\ \mu = \frac{1}{\sum_{n=1}^{2^N}\prod_{k=1}^{L}(\omega_k\beta_{n,k} + 1 - \omega_k\sum_{j=1}^{2^N}\beta_{j,k}) - (2^N - 1)\prod_{k=1}^{L}(1 - \omega_k\sum_{j=1}^{2^N}\beta_{j,k})} \end{gathered} \tag{15}$$

where $\beta_n (n = 1,2,\ldots,2^N)$ denotes the confidence degree of the nth education quality evaluation rating S_n, which satisfies $0 \le \beta_n \le 1$ and $\sum_{n=1}^{2^N} \beta_n = 1$. $\beta_{j,k}$ is the confidence degree of the jth evaluation rating in the posterior of the kth rule.

The output of the EPBRB is expressed as Equation (16):

$$S(y_i)=\{(S_n,\beta_n)\mid n=1,2,\ldots,2^N\} \tag{16}$$

where y_i is the ith attribute corresponding to the input data. $S(\cdot)$ is the inference process of the PBRB.

The final result of the EPBRB model is represented by Equation (17):

$$u(S(y))=\sum_{n=1}^{2^N}u(S_n)\beta_n \tag{17}$$

where $u(S(y))$ denotes the final result of the education quality evaluation. $u(S_n)$ denotes the utility value of the nth evaluation rating S_n.

3.3 Optimization process of the evaluation model based on EPBRB

In the construction process of the EPBRB model, the initial parameters of the ER algorithm and PBRB model are combined with expert knowledge. Due to some ambiguity in expert knowledge, the accuracy of the initial model is poor. To reduce the influence of subjective factors on the evaluation results, the parameters of the two parts are optimized simultaneously to improve the model accuracy with the goal of minimizing the error. The parameters to be optimized and their ranges are shown in Equation (18), and the objective function is shown in Equation (19):

$$\min MSE(\mho)$$
$$\begin{cases} 0\le \gamma_{k,n}\le 1, k=1,2,\ldots,P, n=1,2,\ldots,P_i \\ 0\le q_i\le 1, i=1,2,\ldots,J \\ 0\le \theta_k\le 1, k=1,2,\ldots,L \\ 0\le \delta_i\le 1, i=1,2,\ldots,M \\ 0\le \beta_{n,k}\le 1, n=1,2,\ldots,2^N, k=1,2,\ldots,L \\ \sum_{k=1}^{P}\gamma_{k,n}=1, n=1,2,\ldots,P_i \\ \sum_{n=1}^{2^N}\beta_{n,k}=1, k=1,2,\ldots,L \end{cases} \tag{18}$$

$$MSE(\mho)=\frac{1}{S}\sum_{i=1}^{S}(u(y)-u(y))^2 \tag{19}$$

where $\mho=\{q_i,T_i,\theta_k,\delta_i,\beta_{n,k}\}$ denotes the set of parameters optimized in P-CMA-ES, consisting of V and Q. S is the number of data points in the sample. $\hat{u}(y)$ is the evaluation result of the EPBRB, and $u(y)$ is the true value of the sample.

The projected covariance matrix adaptive evolution strategy (P-CMA-ES) is used as the optimization algorithm in this study. P-CMA-ES is suitable as an optimization method for the EPBRB model because of its advantages, such as good results on high-dimensional nonlinear optimization problems, fast convergence and some interpretability [28]. For the traditional CMA-ES, the equationally constrained solution cannot be obtained because of the strict constraints on the probabilities in the EPBRB. In addition, for strongly constrained problems, the feasible domain is much smaller than the solution space, and reasonable constraint methods are needed [29]. P-CMA-ES adds a projection operation to CMA-ES, to map the unsatisfactory solutions back to the feasible region so that they can satisfy the constraint. P-CMA-ES controls the evolution of the generation toward the optimal solution by continuously updating the covariance matrix of the offspring population. The flow of the algorithm is shown below:

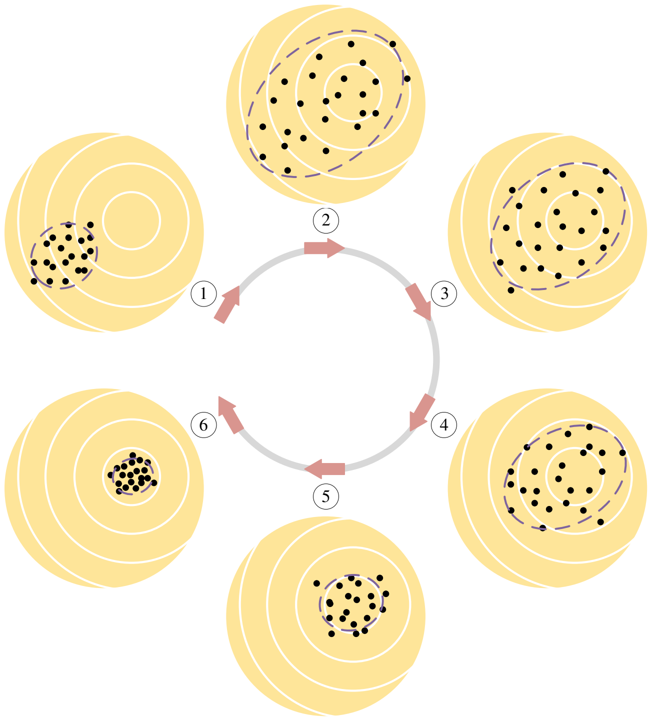

Figure 3. Parameter optimization process based on P-CMA-ES

The detailed steps for P-CMA-ES are listed below:

Step 1: Parameter initialization. Determine the initial parameter set $\omega^0 = \mho^0$, the initial covariance matrix C^0, the initial step size ϵ^0, and the offspring generation size τ.

Step 2: Sampling operation. The population is obtained by Equation (20):

$$\mho_i^{g+1} \sim \omega^g + \varepsilon^g N(0, C^g), i = 1, 2, \dots, \lambda \tag{20}$$

where $\mho_i^{g+1}$ denotes the ith solution of the $(g+1)th$ generation. ω^g denotes the mean of the generation gth. ϵ^g denotes the step size of the gth generation. $N(\cdot)$ denotes the normal distribution. C^g denotes the covariance matrix of the gth generation.

Step 3: Projection operation. The solutions that cannot satisfy the constraint are projected into the feasible region, which is the hyperplane that satisfies the equality constraint. The hyperplane satisfying the equation constraints of the EPBRB model is represented by Equation (21). The projection operation is implemented according to Equation (22):

$$I_e \mho_i^g (1 + n_e \times (j-1) : n_e \times j) = 1 \tag{21}$$

$$\begin{aligned} \mho_i^{g+1}(1 + n_e \times (j-1) : n_e \times j) &= \mho_i^{g+1}(1 + n_e \times (j-1) : n_e \times j) \\ &- \frac{I_e^T}{I_e \times I_e^T} \times \mho_i^{g+1}(1 + n_e \times (j-1) : n_e \times j) \times I_e \end{aligned} \tag{22}$$

where $I_e = [1 \dots 1]_{1\times 2^N}$ denotes the parameter vector. $n_e = 1, \dots, 2^N$ denotes the number of constrained variables. $j = 1, \dots, 2^N + 1$ denotes the number of equality constraints.

Step 4: Selection and Update. τ optimal solutions are selected from the population as the next generation, and their means are updated by the following:

$$\omega^{g+1} = \sum_{i=1}^{\tau} h_i \upsilon_{i:\lambda}^{g+1} \tag{23}$$

where h_i is the weight coefficient. $\upsilon_{i:\lambda}^{g+1}$ denotes the ith solution from λ solutions of the $(g+1)th$ generation.

Step 5: Adaptive operation. By controlling the covariance matrix, the population evolves toward the optimum direction. The updating process is shown as Equation (24):

$$\begin{aligned} C^{g+1} &= (1-c_1-c_2)C^g + c_1 p_c^{g+1}(p_c^{g+1})^T \\ &+ c_1 \sum_{i=1}^{\tau} h_i \left(\frac{(\upsilon_{i:\lambda}^{g+1}-\omega^g)}{\varepsilon^g}\right)\left(\frac{(\upsilon_{i:\lambda}^{g+1}-\omega^g)}{\varepsilon^g}\right)^T \end{aligned} \tag{24}$$

where c_1 and c_2 denote the learning rate. p_c^{g+1} denotes the evolution path of the $(g+1)th$ generation.

Step 6: Repeat the above steps until the optimal solution is obtained.

3.4 Structure of education quality evaluation model based on EPBRB

The modeling process of the education quality evaluation model based on EPBRB is shown in Figure 4. The specific steps are described as follows:

Step 1: Input data processing. The attribute inputs are mapped into a belief distribution about the reference grades of the same attribute using a transformation matrix, and then the ER algorithm is used to fuse the indicators to reasonably reduce the number of attributes in the PBRB.

Step 2: Model construction. The initial parameters are determined based on expert knowledge, the identification framework is extended, and then the education quality evaluation model based on EPBRB is constructed.

Step 3: Parameter optimization. Because the subjective estimation of the initial parameters is not accurate enough, the P-CMA-ES algorithm is used to optimize the model. The optimization goal is to obtain the optimal parameters and improve the accuracy of the evaluation results.

Step 4: Experimental verification. After obtaining the optimized model, experiments are conducted on a real dataset to verify the model validity.

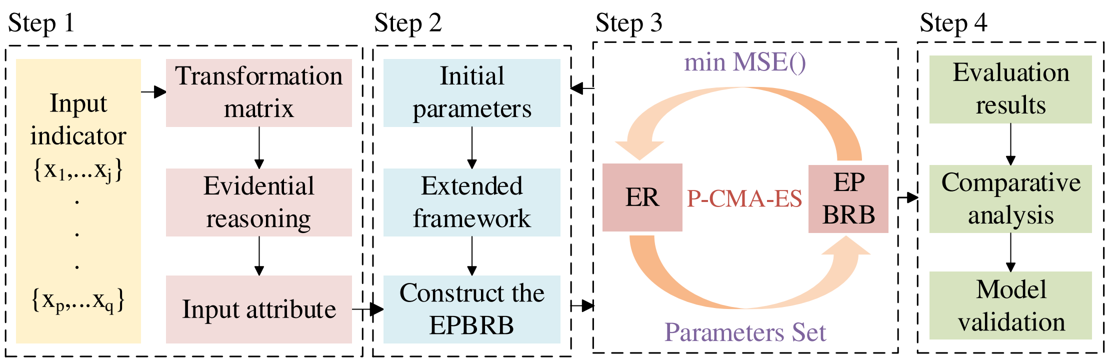

Figure 4. Modeling process of the education quality evaluation model based on the EPBRB

4. Case Study

In this section, the dimensions of education quality evaluation and evaluation indicators are selected based on an actual research study, and the evaluation system is constructed. Then, the model proposed in this paper is tested on the dataset. Finally, the validity of the model is verified through comparative experiments and error analysis. In section 4.1, the data from the experiment are described. In Section 4.2, a model for education quality evaluation is constructed. The model is optimized and tested in Section 4.3. In Section 4.4, a series of comparative experiments are designed, and the results are analyzed and summarized.

4.1 Description of sample data

In the study, data are selected from the Singapore region in TIMSS2019. TIMSS2019 is the seventh international assessment program initiated by the IEA, which uses questionnaires to conduct research in terms of home context, school context, student achievement, teacher and student context, and student-teacher linkages. After analyzing the research of [30–33] et al., this paper selects five dimensions to evaluate the education quality according to the questionnaire content: teacher quality, teaching attitude, teaching content, teaching method and teaching effect. Three or four evaluation indicators are selected under each dimension, and an education quality evaluation system with two levels is constructed, as shown in Table 1.

Table 1. Education quality evaluation system

Index code	First level evaluation indicators	Index code	Secondary evaluation indicators
Y1	Teacher quality	X1	Clear teaching goals
		X2	Solid professional knowledge and accurate problem description
		X3	Detailed and sufficient explanations
		X4	Reasonable course design
Y2	Teaching attitude	X5	Patiently answer questions
		X6	Teach rigorously and conscientiously
		X7	Carefully listen to students' opinions
Y3	Teaching content	X8	Correct concepts and theories
		X9	Extensive and in-depth content
		X10	Combine content with practical applications
Y4	Teaching method	X11	Inspire and encourage students
		X12	Teach using multiple methods
		X13	Teach in line with the student's ability
Y5	Teaching effect	X14	Improved learning ability and interest
		X15	Basic knowledge of students being upgraded
		X16	Upgraded intellectual and moral quality

4.2 Construction of education quality evaluation model based on EPBRB

In the selected dataset, there are 16 input indicators. If these indicators are used as attributes directly, there will be 3^{16} belief rules in the PBRB. For this reason, the secondary indicators are fused by the ER algorithm to the values of the first-level indicators. Taking teacher quality in Table 1 as an example, first, the reference grades of $(x_1 - x_4)$ and y_1 are set. Because they are not consistent, the mapping relationship is constructed using the transformation matrix. Then, the belief distribution of this set of secondary indicators is converted into a belief distribution for y_1. Finally, the utility value of y_1 is calculated based on the fusion result. In this way, 16 secondary indicators are fused into 5 first-level indicators, resulting in a reduction in the attributes of the PBRB model.

The data of input indicator $(x_1 - x_{16})$ were obtained in the form of a questionnaire, pooled in a scale. Based on the available options given in the TIMSS2019 questionnaire, 4 reference grades are designed for describing the endorsement of the secondary indicators, namely, $\{G_1, G_2, G_3, G_4\}$ = {very low (VL), low (L), high (H), very high (VH)}. For each first-level indicator, set the reference grades as $\{R_1, R_2, R_3\}$ = {Poor (P), Medium (M), Good (G)}. According to Section 3.1, 5 first-level indicators and 5 transformation matrices need to be constructed. From Equation (9), the size of each matrix is $[3 \times 4]$. Each initial transformation matrix is shown in the following equation:

$$\mathrm{T}_1 = \begin{bmatrix} 0.9 & 0.6 & 0 & 0 \\ 0.1 & 0.4 & 0.7 & 0 \\ 0 & 0 & 0.3 & 1 \end{bmatrix} \mathrm{T}_2 = \begin{bmatrix} 0 & 0.1 & 0 & 0 \\ 1 & 0.7 & 0.4 & 0.1 \\ 0 & 0.2 & 0.6 & 0.9 \end{bmatrix} \mathrm{T}_3 = \begin{bmatrix} 0.8 & 0.5 & 0 & 0 \\ 0.2 & 0.5 & 0.7 & 0.1 \\ 0 & 0 & 0.3 & 0.9 \end{bmatrix}$$
$$\mathrm{T}_4 = \begin{bmatrix} 0.9 & 0.6 & 0.2 & 0 \\ 0.1 & 0.4 & 0.7 & 0.2 \\ 0 & 0 & 0.1 & 0.8 \end{bmatrix} \mathrm{T}_5 = \begin{bmatrix} 0.7 & 0.4 & 0 & 0 \\ 0.3 & 0.5 & 0.8 & 0.3 \\ 0 & 0.1 & 0.2 & 0.7 \end{bmatrix} \tag{25}$$

The input is transformed into a set of belief degrees of $\{G_1, G_2, G_3, G_4\}$, and then the left-multiplicative transformation matrix is used to consider the resulting belief distribution for $\{R_1, R_2, R_3\}$ as evidence. The initial weight of each piece of evidence is set to 1, and then the input y_i of the PBRB model attributes is obtained by ER fusion and utility calculation. For all five dimensions of $y_1 - y_5$ that are directly involved in the evaluation, three reference points are set, which are ignored (IG), maintained (MA), and emphasized (EM). The initial attribute weights and reference values are shown in Table 2. After indicator fusion, the PBRB has $3^5 = 243$ rules, whose number is significantly reduced, and the initial rule weight is set as 1. The results of the education quality evaluation generated by the model are set to three reference points: low level (LL), average level (AL), and high level (HL). After extension, the identification framework of EPBRB is a power set, and the utility of each rating is shown in Table 3.

Table 2. Reference points, reference values and weights of antecedent attributes

Indicator	Reference point	Reference value	Attribute weight
y_1	(IG,MA,EM)	(3,7,10)	1
y_2	(IG,MA,EM)	(3,7,10)	1
y_3	(IG,MA,EM)	(3,6,10)	1
y_4	(IG,MA,EM)	(3,5,9)	1
y_5	(IG,MA,EM)	(3,5,9)	1

Table 3. The utility of the EPBRB results

Evaluate result	$\emptyset$	LL	AL	HL	{LL,AL}	{LL,HL}	{AL,HL}	{Ω}
Reference value	0	4	9	12	7.5	8	10.5	8.3

4.3 Optimization and testing of the EPBRB model

Model construction needs to incorporate subjective information, resulting in ambiguity in the initial parameters, which need to be optimized. In this subsection, some of the data in the sample are used as training samples to optimize and adjust the model parameters, and the remaining part of the data is used for testing.

The parameters to be optimized in this experiment and their constraints are given in Section 3.3, with a total of 2268 parameters. The 380 groups of data were collated from the questionnaire; 260 groups were randomly selected as training samples, and the remaining data were used for testing.

The P-CMA-ES algorithm is used to globally optimize the parameters, and the number of iteration rounds is set to 600. After optimization, the transformation matrix of the EPBRB model is shown in Equation (26), the evidence weights are shown in Table 4, the attribute weights are shown in Table 5, and the belief rules and their weights are shown in Table 6.

$$
\begin{aligned}
T_1 &= \begin{bmatrix} 0.9586 & 0.932 & 0.5281 & 0.0106 \\ 0.0304 & 0.068 & 0.178 & 0.9577 \\ 0.011 & 0 & 0.2939 & 0.0317 \end{bmatrix} \\
T_2 &= \begin{bmatrix} 0.6667 & 0.9493 & 0.2726 & 0.1153 \\ 0.0584 & 0.0337 & 0.6928 & 0.7026 \\ 0.2749 & 0.017 & 0.0346 & 0.1821 \end{bmatrix} \\
T_3 &= \begin{bmatrix} 0.1408 & 0.7887 & 0.7247 & 0.1011 \\ 0.6641 & 0.0404 & 0.2669 & 0.0623 \\ 0.1951 & 0.1709 & 0.0084 & 0.8366 \end{bmatrix} \\
T_4 &= \begin{bmatrix} 0.3371 & 0.0418 & 0.5951 & 0.0878 \\ 0.6123 & 0.3738 & 0.1427 & 0.4702 \\ 0.0506 & 0.5844 & 0.2622 & 0.442 \end{bmatrix} \\
T_5 &= \begin{bmatrix} 0.3792 & 0.0555 & 0.1414 & 0.2132 \\ 0.2022 & 0.2764 & 0.601 & 0.6936 \\ 0.4186 & 0.6681 & 0.2576 & 0.0932 \end{bmatrix}
\end{aligned} \tag{26}
$$

Table 4. Optimized evidence weights

q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8
0.0667	0.1388	0.907	0.0713	0.1114	0.9902	0.05	0.7818
q_9	q_{10}	q_{11}	q_{12}	q_{13}	q_{14}	q_{15}	q_{16}
0.5101	0.1417	0.4881	0.5399	0.2659	0.1109	0.6163	0.5485

Table 5. Optimized attribute weights

y_1	y_2	y_3	y_4	y_5
0.9377	0.9473	0.3725	0.4131	0.2949

Table 6. Optimized rules and weights

Rule	Weight	Antecedent attributes	Belief distribution of 2^{Ω}
1	0.535	P∧P∧P∧P∧P	{0.0586,0.2103,0.0033,0.0276,0.1846,0.4458,0.0665,0.0033}
2	0.5059	P∧P∧P∧P∧M	{0.0064,0.2406,0.1027,0.2062,0.1205,0.2274,0.0236,0.0726}
3	0.8929	P∧P∧P∧P∧H	{0.0399,0.2757,0.0051,0.1226,0.0074,0.1995,0.2566,0.0932}
4	0.1253	P∧P∧P∧M∧P	{0.0183,0.0873,0.0732,0.2301,0.1447,0.1509,0.0766,0.2189}
…	…	…	…
10	0.7542	P∧P∧M∧P∧P	{0.1265,0.0075,0.1511,0.2216,0.2264,0.058,0.0609,0.148}
…	…	…	…
28	0.4797	P∧M∧P∧P∧P	{0.1039,0.0415,0.0432,0.4812,0.1244,0.0865,0.0957,0.0236}

…	…	…	…
82	0.8371	M^P^P^P^P	{0.1886,0.0707,0.1518,0.0231,0.3126,0.1315,0.085,0.0367}
…	…	…	…
109	0.4086	M^M^P^P^P	{0.2843,0.0284,0.0492,0.0906,0.3008,0.0557,0.1447,0.0463}
…	…	…	…
136	0.394	M^H^P^P^P	{0.242,0.0308,0.1268,0.0509,0.0874,0.1618,0.224,0.0763}
…	…	…	…
163	0.4996	H^P^P^P^P	{0.1953,0.3587,0.0675,0.0535,0.0214,0.0328,0.1411,0.1297}
…	…	…	…
190	0.8394	H^M^P^P^P	{0.0733,0.0526,0.0761,0.1595,0.1881,0.0421,0.2794,0.1289}
…	…	…	…
217	0.3609	H^H^P^P^P	{0.0404,0.3575,0.0526,0.0494,0.047,0.0793,0.3325,0.0413}
…	…	…	…
240	0.5515	H^H^H^M^H	{0.0194,0.0899,0.2945,0.2089,0.0379,0.0367,0.211,0.1017}
241	0.4809	H^H^H^H^P	{0.2441,0.094,0.0361,0.1666,0.192,0.0581,0.0971,0.112}
242	0.1942	H^H^H^H^M	{0.1112,0.043,0.0258,0.1468,0.0187,0.2769,0.3609,0.0167}
243	0.5912	H^H^H^H^H	{0.0473,0.048,0.2545,0.1429,0.1999,0.1518,0.0514,0.1042}

The optimized model is tested, and the actual scores of satisfaction with the education quality are compared with the test results of the model. The MSE, RMSE and MAE are calculated, which are used as the evaluation criteria for the performance of the EPBRB model. The parameters in the following equations have the same meaning as those in Equation (19).

$$MSE = \frac{1}{S}\sum_{i=1}^{S}(u(y)-u(y))^2 \tag{27}$$

$$RMSE = \sqrt{\frac{1}{S}\sum_{i=1}^{S}(u(y)-u(y))^2} \tag{28}$$

$$MAE = \frac{1}{S}\sum_{i=1}^{S}\left|u(y)\text{-}u(y)\right| \tag{29}$$

The smaller the value of MSE is, the smaller the average error between the evaluation result and the actual score, so MSE can measure the accuracy of EPBRB. RMSE is the result of squaring the MSE, which is used as a measure of the deviation between the predicted value and the true value. MAE does not compute the squaring of the error, so it is less sensitive to outliers.

After the experimental measurement, the MSE value of the EPBRB model is 0.2042, the RMSE value is 0.4519, and the MAE value is 0.322. The data comparison is shown in Figure 5. According to the figure and the analysis of the indicators, the model error is small, and the prediction results are close to the real value. Therefore the education quality evaluation model based on EPBRB can accurately produce evaluation results.

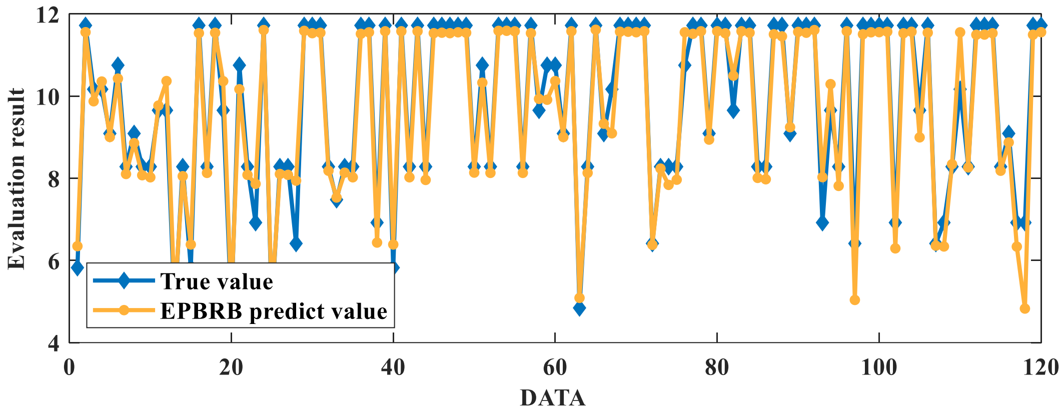

Figure 5. Testing results of EPBRB modeling

4.4 Comparison experiment

To verify the validity of the EPBRB education quality evaluation model, six groups of comparison experiments are designed. First, a comparison is made with the EBRB model based on the original BRB, aiming to verify the influence of the power set identification framework on the accuracy of the evaluation results. Then, a decision tree (DT) and support vector machine (SVM) with certain interpretability are selected for experiments, verifying the superior performance of the EPBRB in terms of interpretability and accuracy. Finally, the evaluation methods are compared with those based on traditional AI algorithms, including the BP neural network (BPNN), k-nearest neighbor (KNN) and random forest (RF). The results of each comparison experiment are shown in Figures 6-11.

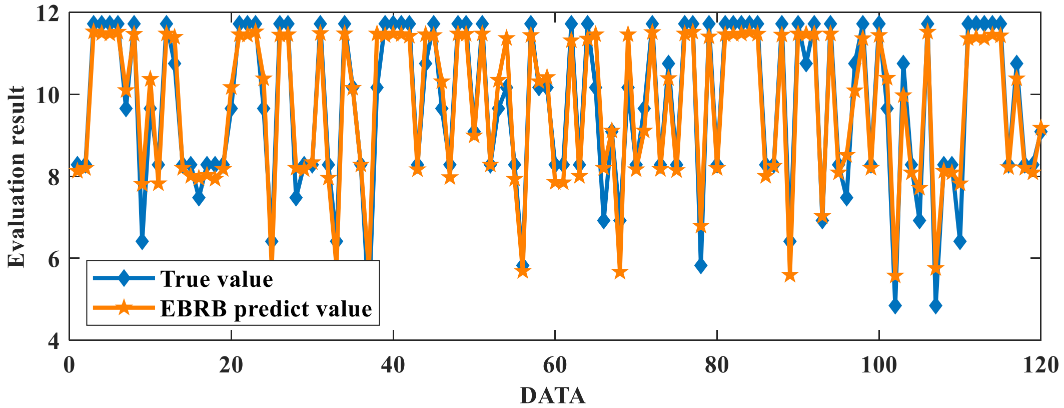

Figure 6. Comparative effect of the EBRB method

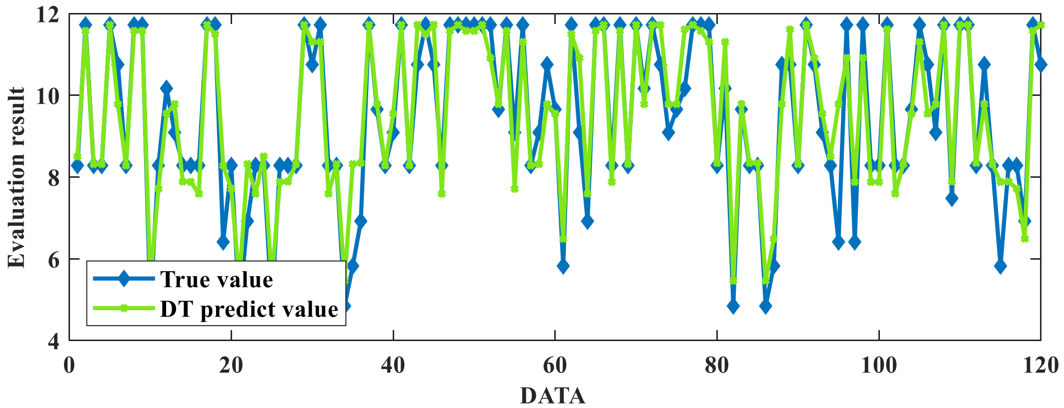

Figure 7. Comparison effect of the DT method

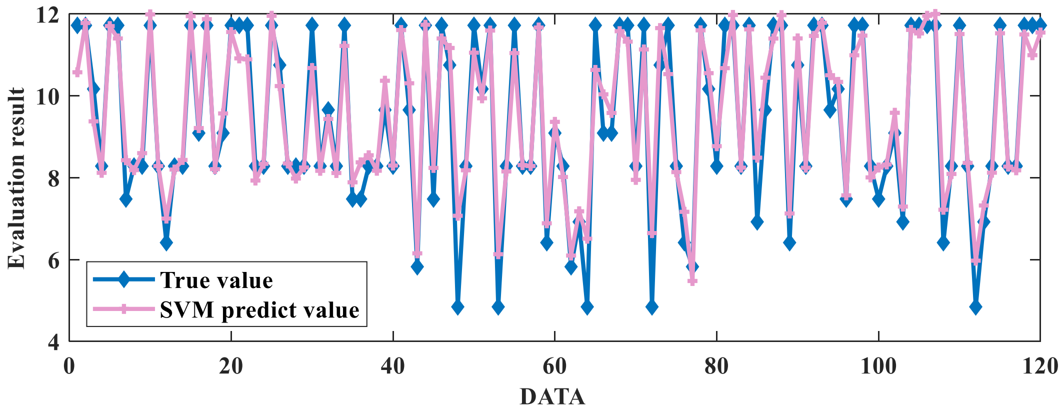

Figure 8. Comparison effect of the SVM method

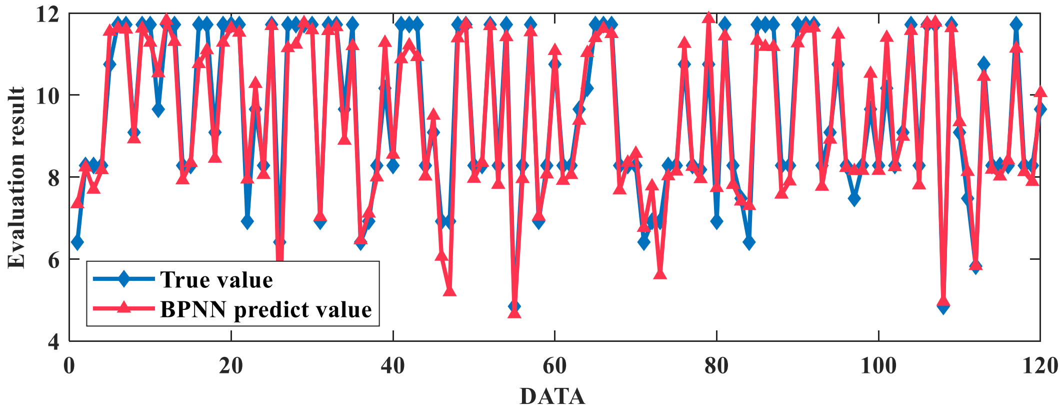

Figure 9. Comparison effect of the BPNN method

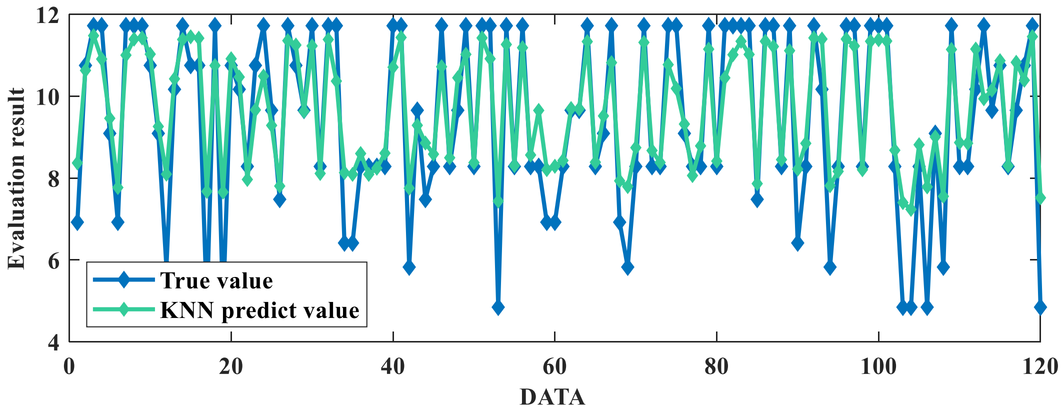

Figure 10. Comparison effect of the KNN method

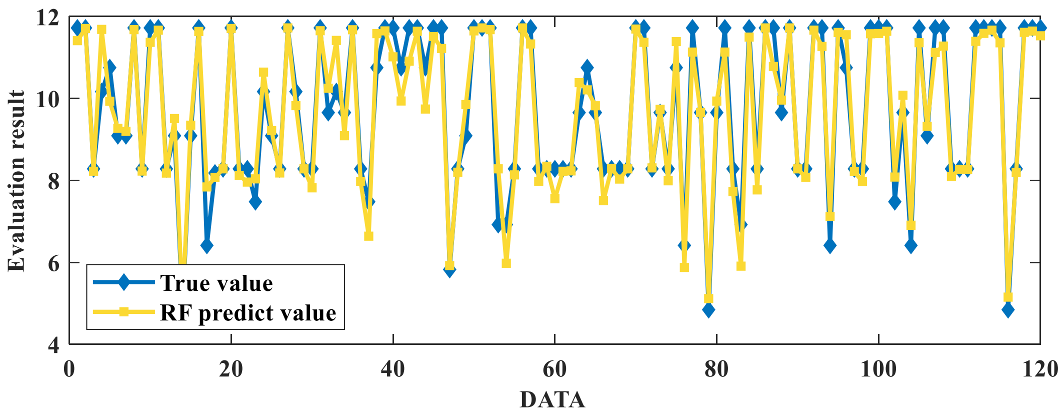

Figure 11. Comparison effect of the RF method

To verify the stability and robustness, 30 experiments are conducted on the EPBRB model and the above six methods, and the average value is taken as the test result. The criterion of each method in each experiment is shown in Figure 12, and the average value of each criterion is shown in Table 7. In addition, to further validate the robustness of the models, repeated experiments are conducted by gradually decreasing the number of training samples. The error criteria of each model under different conditions are shown in Table 8.

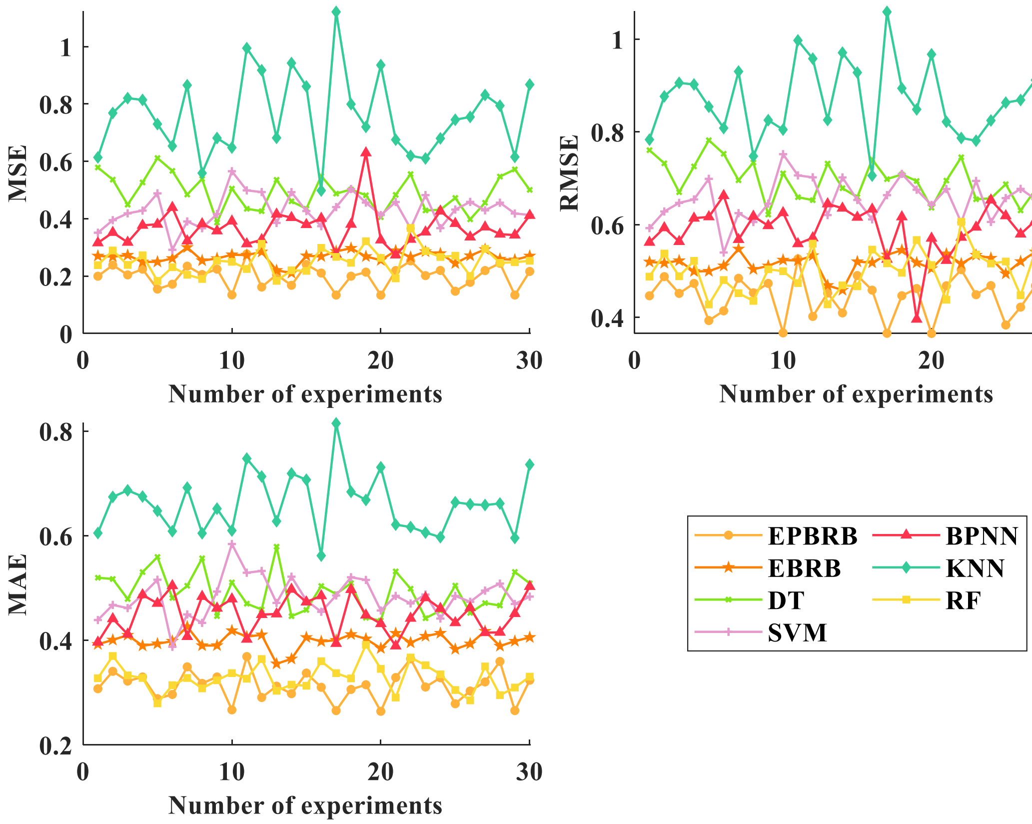

Figure 12. MSE, RMSE and MAE for each experiment

Table 7. Mean values of MSE, RMSE and MAE

Model	EPBRB	EBRB	DT	SVM	BPNN	KNN	RF
Average MSE	0.1989	0.267	0.4915	0.4292	0.3686	0.7607	0.2511
Average RMSE	0.4438	0.5163	0.6998	0.6538	0.5919	0.8687	0.4993
Average MAE	0.3134	0.3989	0.4921	0.4827	0.4508	0.6615	0.3285

Table 8. Comparison of MSE, RMSE and MAE in different experimental conditions

Model	Training:190 Test:190			Training:120 Test:260		
	MSE	RMSE	MAE	MSE	RMSE	MAE
EPBRB	0.2057	0.4535	0.342	0.2094	0.4576	0.3168
DT	0.5113	0.7125	0.4792	0.5289	0.7227	0.4829
SVM	0.4751	0.6878	0.5032	0.5233	0.7205	0.5322
BPNN	0.3994	0.63	0.4859	0.4262	0.6504	0.5057
KNN	0.8268	0.9058	0.6767	0.94	0.9665	0.7621
RF	0.265	0.5137	0.3646	0.2767	0.5238	0.349

According to the data in Table 7, all the criteria of the method based on EPBRB are better than those of the other methods. In the experiments with reduced training samples, the evaluation accuracy is measured by MSE. The accuracy of each method in Table 8 decreased by 3.419%, 4.028%, 10.69%, 8.356%, 8.689% and 5.536% when the training data were 190 groups. The accuracy of each method decreased by 5.279%, 7.609%, 21.92%, 15.63%, 23.57% and 10.2% when the training data were 120 groups.

Compared with EBRB, although the model is also constructed based on BRB, the evaluation results of EPBRB have less error. It can be concluded that local ignorance and global ignorance due to similar input data affect the performance of the model, and extending the identification framework to power sets can improve the accuracy of the BRB. DT and SVM also have some interpretability, where DT relies on tree structure for decision making and SVM classifies according to hyperplane. According to Section 3.2, EPBRB has a rigorous mathematical derivation process, which expresses the likelihood of each evaluation rating in terms of the belief degree, so that EPBBR has a stronger interpretability than DT and SVM and has a higher accuracy. BPNN, KNN and RF are black-box methods. The initial parameters are random, and the derivation process from input to output is completely unknowable. In contrast, EPBRB can combine expert knowledge to construct the initial model and optimize the model parameters with existing data as training samples. Meanwhile, the causal relation of rules in if-then form is explicit, and the method of inference on the basis of belief rules is clearer and more transparent, so the inference results can be traced. As typical data-driven models, SVM, BPNN, KNN and RF are simple to model and can directly accept data to establish nonlinear relationships. However, these types of methods rely on the amount of data, and when the training samples are insufficient, the model effect will be significantly reduced.

Through the above experiments and analysis, the education quality evaluation model based on the EPBRB has the following advantages. First, EPBRB can deal with the ignorance problem. By assigning a belief degree to the power set identification framework, various results can be expressed, so the evaluation results are more accurate. Second the construction process of the EPBRB can be combined with expert knowledge, the inference process is completely transparent, and the evaluation results are more credible. Finally, the EPBRB model has a high learning efficiency, can better results under the condition of small samples, and has better generalizability.

5. Conclusions

Education quality evaluation is a direct means of promoting education development, and many evaluation methods based on information technology have been proposed, but there are still some problems that deserve attention. First, the results of education quality evaluation are very sensitive, and results with insufficient credibility may stimulate educational imbalance. Second, most of the education-related data come from questionnaires, but the effective return rate of questionnaires is low, and the amount of available data is not large. Finally, the model may not give clear evaluation results when the input indicators are similar.

Considering the above problems, this paper proposes a new education quality evaluation model based on EPBRB. The method improves the BRB model with strong interpretability. For model inputs, the transformation matrices are used to convert different indicators into belief distributions under the same attribute reference grades, and then the inputs are obtained through ER fusion as a way to reasonably reduce the number of attributes in the BRB. For the model output, the power set of evaluation results is used as the identification framework so that the model can express local ignorance and global ignorance. Finally, the initial parameters are optimized using P-CMA-ES to improve the model accuracy.

In the case study part of this paper, the stability and robustness of the new method is verified through comparative experiments and error analysis. The experimental results show that the evaluation results of the EPBRB model are more accurate and that the model still shows better results under the condition of less available data. However, there are still aspects to be improved in this study. The most important of these is to select evaluation indicators of education quality from a more comprehensive dimension and to screen and verify them to construct a more scientific evaluation system of education quality. Next, a sensitivity analysis method is designed to find the factors that have a greater impact on education quality and to analyze and estimate the degree of their impact and sensitivity on education quality so that educators can clarify the weak points in teaching.

Data Availability Statement

Data for this study were taken from https://timssandpirls.bc.edu/databases-landing.html.

Declaration of interest statement

The authors declare no conflict of interest.

References

[1] S. Liu, Research on the teaching quality evaluation of physical education with intuitionistic fuzzy TOPSIS method, J. Intell. Fuzzy Syst. 40 (2021) 9227–9236. https://doi.org/10.3233/JIFS-201672.

[2] J. Hou, Online teaching quality evaluation model based on support vector machine and decision tree, J. Intell. Fuzzy Syst. 40 (2021) 2193–2203. https://doi.org/10.3233/JIFS-189218.

[3] H.-X. Liu, Y.-H. Zhang, S.-B. Tsai, Cloud Education Chain and Education Quality Evaluation Based on Hybrid Quantum Neural Network Algorithm, Wirel. Commun. Mob. Comput. 2021 (2021) 1909345. https://doi.org/10.1155/2021/1909345.

[4] X. Fan, S. Tian, Z. Lu, Y. Cao, Quality evaluation of entrepreneurship education in higher education based on CIPP model and AHP-FCE methods, Front. Psychol. 13 (2022) 973511. https://doi.org/10.3389/fpsyg.2022.973511.

[5] M. Li, Y. Su, Evaluation of Online Teaching Quality of Basic Education Based on Artificial Intelligence, Int. J. Emerg. Technol. Learn. 15 (2020) 147–161. https://doi.org/10.3991/ijet.v15i16.15937.

[6] X. Lv, A Quality Evaluation Scheme for Curriculum in Ideological and Political Education Based on Data Mining, in: 2021 13th International Conference on Measuring Technology and Mechatronics Automation (Icmtma 2021), Ieee Computer Soc, Los Alamitos, 2021: pp. 649–652. https://doi.org/10.1109/ICMTMA52658.2021.00149.

[7] H. Wang, Teaching quality monitoring and evaluation using 6G internet of things communication and data mining, Int. J. Syst. Assur. Eng. Manag. 14 (2023) 120–127. https://doi.org/10.1007/s13198-021-01206-8.

[8] C. Liu, Y. Feng, Y. Wang, An innovative evaluation method for undergraduate education: an approach based on BP neural network and stress testing, Stud. High. Educ. 47 (2022) 212–228. https://doi.org/10.1080/03075079.2020.1739013.

[9] P. Niu, An artificial intelligence method for comprehensive evaluation of preschool

education quality, Front. Psychol. 13 (2022) 955870. https://doi.org/10.3389/fpsyg.2022.955870.
[10] L. Fang, Construction of Physical Education Quality Evaluation Index and Analysis with Wearable Device, Comput. Intell. Neurosci. 2022 (2022) 1190394. https://doi.org/10.1155/2022/1190394.
[11] T. Yuan, Algorithm of Classroom Teaching Quality Evaluation Based on Markov Chain, Complexity. 2021 (2021) 9943865. https://doi.org/10.1155/2021/9943865.
[12] D. Fouskakis, G. Petrakos, I. Vavouras, A Bayesian hierarchical model for comparative evaluation of teaching quality indicators in higher education, J. Appl. Stat. 43 (2016) 195–211. https://doi.org/10.1080/02664763.2015.1054793.
[13] X. Zhao, C. Zheng, Fuzzy Evaluation of Physical Education Teaching Quality in Colleges Based on Analytic Hierarchy Process, Int. J. Emerg. Technol. Learn. 16 (2021) 217–230. https://doi.org/10.3991/ijet.v16i06.21097.
[14] G. Ning, Evaluation Criteria for Quality Education of Physical Education Lessons Based on Logical Analysis, Int. J. Emerg. Technol. Learn. 16 (2021) 87–99. https://doi.org/10.3991/ijet.v16i21.26867.
[15] J.V. Tu, Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes, Journal of Clinical Epidemiology. 49 (1996) 1225–1231. https://doi.org/10.1016/S0895-4356(96)00002-9.
[16] Jian-Bo Yang, Jun Liu, Jin Wang, How-Sing Sii, Hong-Wei Wang, Belief rule-base inference methodology using the evidential reasoning Approach-RIMER, IEEE Trans. Syst., Man, Cybern. A. 36 (2006) 266–285. https://doi.org/10.1109/TSMCA.2005.851270.
[17] Z.-J. Zhou, G.-Y. Hu, C.-H. Hu, C.-L. Wen, L.-L. Chang, A Survey of Belief Rule-Base Expert System, IEEE Trans. Syst. Man Cybern, Syst. 51 (2021) 4944–4958. https://doi.org/10.1109/TSMC.2019.2944893.
[18] B. Zhang, Y. Zhang, G. Hu, Z. Zhou, L. Wu, S. Lv, A method of automatically generating initial parameters for large-scale belief rule base, Knowledge-Based Systems. 199 (2020) 105904. https://doi.org/10.1016/j.knosys.2020.105904.
[19] L. Chang, L. Zhang, X. Xu, Correlation-oriented complex system structural risk assessment using Copula and belief rule base, Information Sciences. 564 (2021) 220–236. https://doi.org/10.1016/j.ins.2021.02.076.
[20] Q. Jia, J. Hu, W. Zhang, A fault detection method for FADS system based on interval-valued neutrosophic sets, belief rule base, and D-S evidence reasoning, Aerospace Science and Technology. 114 (2021) 106758. https://doi.org/10.1016/j.ast.2021.106758.
[21] D. Tang, J.-B. Yang, K.-S. Chin, Z.S.Y. Wong, X. Liu, A methodology to generate a belief rule base for customer perception risk analysis in new product development, Expert Systems with Applications. 38 (2011) 5373–5383. https://doi.org/10.1016/j.eswa.2010.10.018.
[22] X. Xu, H. Xu, C. Wen, J. Li, P. Hou, J. Zhang, A belief rule-based evidence updating method for industrial alarm system design, Control Engineering Practice. 81 (2018) 73–84. https://doi.org/10.1016/j.conengprac.2018.09.001.
[23] W. Han, X. Kang, W. He, L. Jiang, H. Li, B. Xu, A new method for disease diagnosis based on hierarchical BRB with power set, Heliyon. 9 (2023) e13619. https://doi.org/10.1016/j.heliyon.2023.e13619.

[24] Y. Xie, W. He, H. Zhu, R. Yang, Q. Mu, A new unmanned aerial vehicle intrusion detection method based on belief rule base with evidential reasoning, Heliyon. 8 (2022) e10481. https://doi.org/10.1016/j.heliyon.2022.e10481.

[25] Z. Zhou, Z. Feng, C. Hu, X. Han, Z. Zhou, G. Li, A hidden fault prediction model based on the belief rule base with power set and considering attribute reliability, Sci. China Inf. Sci. 62 (2019) 202202. https://doi.org/10.1007/s11432-018-9620-7.

[26] L. Chang, Y. Chen, Z. Hao, Z. Zhou, X. Xu, X. Tan, Indirect disjunctive belief rule base modeling using limited conjunctive rules: Two possible means, International Journal of Approximate Reasoning. 108 (2019) 1–20. https://doi.org/10.1016/j.ijar.2019.02.006.

[27] H. Chen, G. Zhou, X. Zhang, H. Zhu, W. He, Learning Emotion Assessment Method Based on Belief Rule Base and Evidential Reasoning, Mathematics. 11 (2023) 1152. https://doi.org/10.3390/math11051152.

[28] Y. Cao, Z. Zhou, C. Hu, W. He, S. Tang, On the Interpretability of Belief Rule-Based Expert Systems, IEEE Transactions on Fuzzy Systems. 29 (2021) 3489–3503. https://doi.org/10.1109/TFUZZ.2020.3024024.

[29] Z.-J. Zhou, G.-Y. Hu, B.-C. Zhang, C.-H. Hu, Z.-G. Zhou, P.-L. Qiao, A Model for Hidden Behavior Prediction of Complex Systems Based on Belief Rule Base and Power Set, IEEE Trans. Syst. Man Cybern, Syst. 48 (2018) 1649–1655. https://doi.org/10.1109/TSMC.2017.2665880.

[30] H. Wang, F. Yang, X. Xing, Evaluation Method of Physical Education Teaching and Training Quality Based on Deep Learning, Comput. Intell. Neurosci. 2022 (2022) 1680888. https://doi.org/10.1155/2022/1680888.

[31] D. Zhang, L. Chen, A BP Neural Network-Assisted Smart Decision Method for Education Quality, IEEE Access. 11 (2023) 74569–74578. https://doi.org/10.1109/ACCESS.2023.3294804.

[32] L. Kong, X. Cai, H. Lu, Quality evaluation system of engineering cost education curriculum based on data clustering, EAI Endorsed Trans. Scalable Inform. Syst. (2022). https://doi.org/10.4108/eai.11-2-2022.173451.

[33] H. Di, H. Zhang, P. Li, Teaching Quality of Ideological and Political Education in Colleges Based on Deep Learning, Int. J. e-Collab. 19 (2023) 18–18. https://doi.org/10.4018/IJeC.316829.