Article

Learning Emotion Assessment Method Based on Belief Rule Base and Evidential Reasoning

Haobing Chen [1], Guohui Zhou [1,*], Xin Zhang [2,*], Hailong Zhu [1] and Wei He [1,2]

1 School of Computer Science and Information Engineering, Harbin Normal University, Harbin 150025, China
2 High-Tech Institute of Xi'an, Xi'an 710025, China
* Correspondence: zhouguohui@hrbnu.edu.cn (G.Z.); zhangxin_shanxi@163.com (X.Z.)

Abstract: Learning emotion assessment is a non-negligible step in analyzing learners' cognitive processing. Data are the basis of the learning emotion assessment. However, the existing learning emotion assessment models cannot balance model accuracy and interpretability well due to the influence of uncertainty in the process of data collection and model parameter errors. Given the above problems, a new learning emotion assessment model based on evidence reasoning and a belief rule base (E-BRB) is proposed in this paper. First, the transformation matrix is introduced to transform multiple emotional indicators into the same standard framework and integrate them, which keeps the consistency of information transformation. Second, the relationship between emotional indicators and learning emotion states is modeled by E-BRB in conjunction with expert knowledge. In addition, we employ the projection covariance matrix adaptation evolution strategy (P-CMA-ES) to optimize the model parameters and improve the model's accuracy. Finally, to demonstrate the effectiveness of the proposed model, it is applied to emotion assessment in science learning. The experimental results show that the model has better accuracy than data-driven models such as neural networks.

1. Introduction

Learning emotion is one of the essential factors affecting cognitive processing and the learning effect [1]. On the one hand, the learners' emotional state can indicate the learners' preferences for teaching content and teaching environment, which helps explore deep cognitive styles and learning interests [2]. On the other hand, it can reflect the influence mechanism of learners' knowledge level and cognitive structure on their subjective learning experience, which helps to reveal the deep learning mechanism [3]. However, unreliable and uninterpretable results can increase the ethical risk of educational practice. Therefore, an accurate and interpretable learning emotional state assessment is of great significance in the context of the rapid development of intelligent technology.

In recent years, learning emotion assessment has progressively attracted the attention of researchers. Several different assessment approaches have been presented, which can be divided into the following three types: (I) Data-driven models, which utilize a large amount of training data to obtain a prediction model. Estrada et al. used three different techniques of machine learning, including support vector machine (SVM) and random forest, deep learning, such as convolutional neural networks (CNNs) and long short-term memory networks (LSTM), and evolutionary algorithms to evaluate students' learning emotions. Furthermore, the three methods were compared and analyzed [4]. Ashwin and Guddeti proposed a novel mixed convolutional neural network architecture for analyzing students' emotional states in a classroom setting. The architecture was divided into two parts. In

terms of individual image frames, the first model aimed to recognize a single student's emotional state, and the second analyzed several students. The whole mixed architecture was used to predict the overall emotional state of the whole class [5]. Bota et al. used multimodal physiological data to assess emotions concerning low/high arousal and valence classification using supervised learning, decision fusion, and feature fusion techniques. The experiment tested seven methods, including K-nearest neighbor, decision tree, random forest; support vector machines, AdaBoost, Gaussian naive Bayes, and quadratic discriminant analysis [6]. Chan et al. used deep learning techniques to analyze multimodal data generated in the learning process. Students' emotional attitudes, academic engagement, and classroom concentration were quantitatively evaluated to analyze learners' learning motivation [7]. (II) Knowledge-driven models, which use expert experience and domain knowledge to establish the relationship between data and emotional states, can provide transparent modeling processes and interpretable results. Hwang et al. proposed an expert system approach that considers individual learners' emotional and cognitive status. The learning system developed was composed of four modules, of which the expert system module uses fuzzy reasoning for analyzing the student's emotional state [8]. Fodor et al. built a sensory network to study physiological data collection and used the collected data for emotional state identification. A Petri net model that simulates how certain emotions affect physiological data was constructed to reduce the invasiveness of data collection [9]. Kurniawan et al. utilized an attitude questionnaire and an interview form to assess Indonesian students' attitudes toward natural sciences. Descriptive statistics were used for the attitude questionnaire, and Miles and Huberman's models were used for the interview data [10]. (III) Hybrid-driven models, such as the hidden Markov model, Bayesian network, and belief function-based model [11], utilize quantitative data and qualitative knowledge to establish the model. Patlar Akbulut presented a method to accurately recognize six emotions using electrocardiogram (ECG) and electrodermal activity (EDA) signals and applying autoregressive hidden Markov models (AR-HMMs) and heart rate variability analysis on these signals [12]. Harper et al. proposed an end-to-end model for classifying emotions from unimodal data. In addition, a Bayesian framework for uncertainty modeling was further proposed. It describes a probabilistic process for accepting or rejecting the model output depending on the intended application [13]. Ray et al. combined deep learning methods with rule-based approaches to improve model performance in terms of aspect extraction and sentiment scoring. On the one hand, a seven-layer specific CNN structure was developed. The concept of rule-based was introduced in order to improve the performance of aspect extraction [14].

Although the above methods can be applied to learning emotion assessment, there are still some problems. The data-driven-based methods rely too much on samples to train the model, which is unsuitable for small sample sizes. With a large sample of data, a data-driven model can be used to construct a more accurate assessment model. However, data-driven models cannot achieve a balance between model accuracy and model interpretability. Models with higher precision may have worse interpretability. From the standpoint of education, the low interpretability of the final result makes it difficult for people to understand the final result, so it is impossible to determine which factor is the dominant factor that activates negative emotions [15]. In addition, the lack of interpretability of algorithms leads to information asymmetry between algorithm developers and users, which will increase the inequity of education. The knowledge-driven method is not conducive to improving evaluation accuracy and is also slightly insufficient in addressing uncertainty. When the data collection technology is not mature, learners are easily affected by environmental, physical, and psychological factors, and the measured values often have significant differences and fluctuations [16]. In this case, it is difficult to obtain accurate results based on knowledge-driven methods. The hybrid-driven method combines domain expert knowledge and experience as well as historical data, which can make the output of the model more accurate. It maintains good performance in terms of model interpretability

and accuracy simultaneously. The belief rule base (BRB) is introduced in this paper to achieve a balance between accuracy and interpretability.

As a gray-box model, the BRB model can express various types of uncertain information, and its reasoning process and output results are transparent and interpretable [17]. BRB provides an information scheme for formulating expert experience, uncertain knowledge, and hybrid information [18]. It has been widely used in fault diagnosis [19], Complex system modeling [20], state assessment [21], and medical science [22]. The emotion assessment model based on BRB is an ideal choice for this situation.

Learning emotion assessment faces many problems, such as many types of indicators and complex relationships. When establishing a model using BRB, it is necessary to traverse all reference values of all antecedent attributes. Therefore, the number of emotional indicators directly influences the complexity and structure of the model [23]. Too many antecedent attributes can lead to combinatorial explosion problems, which restricts the suitability of the BRB model in higher dimensional problems. For the combinatorial explosion problem, commonly used methods include principal component analysis (PCA) [24], rough set theory [25], gray target (GT) [26], etc. However, these methods may lose some information and reduce the model's accuracy when there are no significant variations in the degree to which the different attributes affect the consequent parts. In this paper, the complexity of the indicators is reduced by fusing multiple learning sentiment indicators. As an information fusion mechanism, the evidential reasoning (ER) algorithm can avoid information loss, give reliable fusion results, and achieve effective data analysis [27]. The fusion of learning emotional indicators through the ER algorithm can effectively avoid the combinatorial explosion problem. At the same time, the model adopts global optimization to prevent the overall model from falling into a local optimum that affects the performance of the model.

Therefore, a learning emotion assessment model based on evidential reasoning and belief rule base (E-BRB) is proposed in this paper. Firstly, multiple emotional indicators are converted to belief distribution under the predetermined framework by the transformation matrix. Then, the ER algorithm is used to fuse the information of similar learning emotion indicators, and the results are used as input to establish the E-BRB model. Finally, the optimal model is obtained by the optimization algorithm. The main contributions of this paper are as follows:

(1) On the basis of the transformation matrix, the mapping relationship between learning emotion indicators and fusion results is built, which solves the problem of inconsistent emotion indicators reference grades and result grades in educational practice, ensures the integrity of information transformation, and avoids the loss of information.

(2) A learning emotion assessment model based on E-BRB is constructed. The model solves the combinatorial explosion problem of BRB by using the practical information fusion ability and efficient reasoning ability of E-BRB. Thus, the learning emotion assessment in an educational environment is achieved. At the same time, the model considers both accuracy and interpretability, reducing the potential and ethical risks of educational decision-making.

The structure of this paper is as follows. In Section 2, two problems in learning emotion assessment and their solutions are analyzed. In Section 3, the learning emotion assessment model based on E-BRB is established. In Section 4, an experimental case study is designed to verify the validity of the E-BRB model. In Section 5, our conclusion is summarized, and future work for learning emotion assessment is discussed.

2. Problem Formulation

The problems in the assessment of learning emotion are described and analyzed in Section 2.1. Aiming at the existing problems, we propose a learning emotion assessment model based on E-BRB in Section 2.2.

2.1. Problem Formulation of Learning Emotion Assessment

Many methods and approaches have been utilized for learning emotion assessment. Traditional learning emotion assessment models cannot address complex indicators well, and most methods lack transparency and interpretability. In most cases, we pay more attention to why and how the model obtains results. Therefore, it is of great significance to construct an accurate learning emotion assessment model in an interpretable way. The problems existing in the assessment of learning emotion are described, and a learning emotion evaluation model based on E-BRB is proposed in this section. This paper mainly focuses on the following three problems:

Problem 1. *The first problem to be solved is how to establish the mapping relation between the reference grade of the emotion indicators and the result grade. The evaluation indicators of learning emotion include learning interest, learning attitude, learning will, academic values, learning motivation, and learning beliefs. There are many types of indicators and complex relationships that need to be integrated at different levels. The relationship between the emotion indicator reference grade and result grades may not correspond in the indicator fusion process. The accuracy of the output results may be affected if the relationship between the reference grade of the input indicators and the result grade is forced to be considered as a one-to-one correspondence. The mapping relationship is established based on the transformation matrix, and an input information transformation framework is built. The mapping relationship is shown in Equation (1):*

$$(F_1, F_2, \ldots, F_P) = \mathbf{Z}_i(H_{1,i}, H_{2,i}, \ldots, H_{P_i,i}) \tag{1}$$

where $(F_1, F_2, \ldots, F_P)$ denotes P result grades; $\mathbf{Z}_i$ represents the mapping function between the ith input indicator reference grade and the result state; and $(H_{1,i}, H_{2,i}, \ldots, H_{P_i,i})$ denotes ith input indicator P_i reference grades.

Problem 2. *How to reasonably construct and optimize the learning emotion assessment model is the second problem to be solved in this paper. In the process of learning emotion assessment, the interpretability of the model is an essential reference factor for modeling. As a gray-box model, BRB has superior non-linear modeling capability while ensuring model interpretability. During BRB construction, each possible combination of all reference values for all attributes needs to be covered. When there are too many input attributes, it will lead to the BRB combinatorial explosion problem. The emotion assessment model based on E-BRB is proposed in this paper. Multiple attributes are fused using the ER algorithm, and the fusion results are fed into the BRB model. The parameters in the model are determined by the incomplete knowledge of experts. The initial parameters of the model may not be accurate due to the ambiguity of the knowledge representation, so an optimization algorithm must be utilized to optimize the parameters. In this paper, an optimization model was constructed based on the projection covariance matrix adaptation evolution strategy (P-CMA-ES). The construction process of the model can be described in Equations (2) and (3). The optimization process of the model is shown in Equation (4):*

$$y_i = ER(x_1, x_2, \cdots, x_J, \alpha) \tag{2}$$

$$u(S(y)) = EBRB(y_1, y_2, \cdots, y_M, \beta) \tag{3}$$

$$\Omega = \psi(EBRB(\cdot)) \tag{4}$$

where the fusion process of multiple attributes is shown in Equation (2); y_i is the result obtained after fusion of the ER algorithm; $ER(\cdot)$ represents fusion functions; $x_1, x_2, \cdots, x_J$ denote J attributes; α is the vector of parameters in the fusion process; $u(\cdot)$ denotes the result of learning emotion assessment; $S(\cdot)$ denotes the learning emotion grade; $EBRB(\cdot)$ represents the reasoning process of the model; and β represents the set of parameters in the model reasoning process; Ω represents the set of parameters that need to be optimized in the E-BRB model and $\psi(\cdot)$ denotes the model optimization process using the P-CMA-ES optimization algorithm.

2.2. Construction of the New Learning Emotion Assessment Model

In response to the above two problems, we propose a learning emotion assessment model based on E-BRB in this paper, which contains L belief rules. Assuming that $x_i = 1, \ldots, J$ are the learning emotion indicators, $y_i, i = 1, \ldots, M$ are obtained after ER algorithm fusion. The kth rule in the E-BRB model can be described as:

$$
\begin{aligned}
R_k : \ & IF \left(y_1 \ is \ A_1^k \right) \wedge \left(y_2 \ is \ A_2^k \right) \wedge \cdots \wedge \left(y_M \ is \ A_M^k \right) \\
& THEN \left\{ (D_1, \beta_{1,k}), (D_2, \beta_{2,k}), \ldots, (D_N, \beta_{N,k}) \right\}, \left(\sum_{n=1}^{N} \beta_{n,k} \leq 1 \right) \\
& with \ a \ rule \ weight \ \theta_k (k = 1, 2, \ldots, L) \\
& and \ attribute \ weights \ \delta_1, \delta_2, \ldots \delta_M
\end{aligned}
\tag{5}
$$

where $R_k(k = 1, 2, \ldots, L)$ is the kth rule in the E-BRB model; $A_i^k (i = 1, \ldots, M)$ represents the reference value of the ith antecedent attribute in the kth rule; M denotes the number of antecedent attributes in the kth rule; and $(y_1, y_2, \ldots y_M)$ is the feature that can reflect the emotional state of students, which is the result of the fusion of the ER algorithm. The number of antecedent attributes in this model depends on the number of results after ER algorithm fusion. When the input vector of the kth rule satisfies $(y_1, y_2, \ldots, y_M) = (A_1^k, A_2^k, \ldots, A_M^k)$, the belief degree corresponding to the emotional state D_n is $\beta_{n,k}(n = 1, 2, \ldots, N)$. $\sum_{n=1}^{N} \beta_{n,k} \leq 1$, if $\sum_{n=1}^{N} \beta_{n,k} = 1$, the kth rule is said to be complete; otherwise, it is incomplete. L denotes the total number of rules in the E-BRB. θ_k denotes the weight of the kth rule. $\delta_i(i = 1, 2, \ldots M)$ denotes the weight of the ith antecedent attribute.

Remark 1. *The E-BRB model has two parts: the ER algorithm and the classic BRB model. First, the ER algorithm is applied to integrate similar indicators, and then the output of the ER algorithm is used as the input of the classical BRB model.*

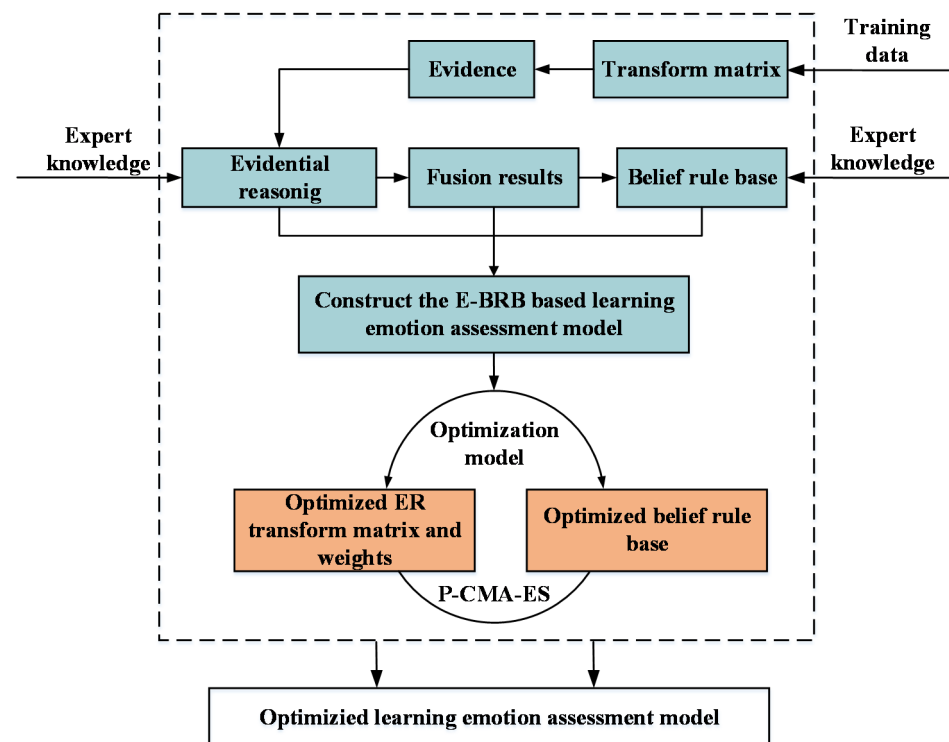The modeling process of the model is shown in Figure 1.

Figure 1. The modeling process of the E-BRB model.

3. Learning Emotion Assessment Model Based on E-BRB

A learning emotion assessment model based on E-BRB is proposed to address the three problems mentioned in Section 2.1. A transformation matrix for addressing inconsistent input–output mappings is presented in Section 3.1. Then, the inference of the E-BRB model is described in Section 3.2. An optimization model is proposed in Section 3.3 to train the parameters in the model, which uses the P-CMA-ES algorithm as the optimization algorithm. A learning emotion assessment modeling method based on the E-BRB model is proposed in Section 3.4.

3.1. Transformation Method of Input Indicators

When using the ER algorithm to fuse learners' emotional indicators, a set of result grades is predetermined, which are mutually exclusive and collectively exhaustive. After determining the emotional indicators, the input indicator reference grade is introduced to obtain the initial evidence pointing to the result grade. The indicator reference grades, as an essential part of the fusion process, significantly impact the belief distribution of the initial evidence. Finally, the initial evidence and the weight of evidence are fused using the ER algorithm to obtain the fusion results. The above process confirms that the reference grade of the input indicator corresponds to the result grade one by one. However, in the actual learning emotion assessment, the emotional state grade is predetermined, resulting in inconsistency with the input indicator reference grade. For example, according to the assessment items, the input reference grade can be easily divided into "enjoyment" and "disgust", but the emotional state grade is preset as "joy", "boredom", and "confusion". A transformation matrix is proposed to address the above problem in this subsection to solve the problem of inconsistency between the reference grades of input indicators and the result grades.

Let us suppose there are J indicators and P emotional state grades, which can be expressed as $\{F_n | n = 1, 2, \ldots, P\}$. For the ith input attribute, the number of input indicator reference grades is P_i, which can be expressed as $\{H_{n,i} | n = 1, 2, \ldots, P_i\}$. $H^i = \{H_{1,i}, H_{2,i}, \ldots, H_{P_i,i}\}$ and $F = \{F_1, F_2, \ldots, F_P\}$ are sets of mutually exclusive and exhaustive propositions. H^i and F are represented as discernment framework 1 and discernment framework 2, respectively. The transformation of the input information is shown in Figure 2. The specific transformation process from discernment framework 1 to discernment framework 2 is as follows:
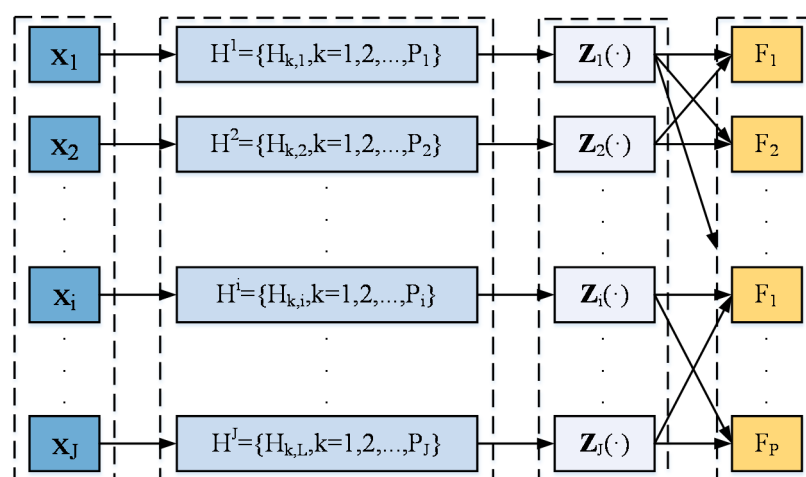

Figure 2. The transformation between input indicator and output.

Firstly, the correspondence between the kth referential grade $H_{k,i}$ of the ith emotion indicator and the emotional state grade $F = \{F_1, F_2, \ldots, F_P\}$ can be described by the "IF-THEN" rule as follows:

$$R_{k,i} : \text{IF } x_i = H_{k,i}, \text{ THEN } \left\{(F_1, z_{1,k}), (F_2, z_{2,k}), \ldots, (F_P, z_{P,k})\right\},$$
$$\left(\sum_{n=1}^{P} z_{n,k} = 1, 0 \leq z_{n,k} \leq 1\right) \tag{6}$$

where $R_{k,i}$ denotes the kth rule for the ith emotion indicator. $z_{n,k}$ denotes the belief degree corresponding to the consequent F_n when the referential grade of the emotion indicator x_i is $H_{k,i}$.

Then, the mapping relationship between discernment framework 1 and discernment framework 2 can be determined by P_i rules, which can be represented by the following matrix:

$$\mathbf{Z}_i = \begin{matrix} & \begin{matrix} H_{1,i} & H_{2,i} & \ldots & H_{P_i,i} \end{matrix} \\ \begin{matrix} F_1 \\ F_2 \\ \vdots \\ F_P \end{matrix} & \begin{bmatrix} z_{1,1} & z_{1,2} & \ldots & z_{1,P_i} \\ z_{2,1} & z_{2,2} & \ldots & z_{2,P_i} \\ \vdots & \vdots & \ddots & \vdots \\ z_{P,1} & z_{P,2} & \ldots & z_{P,P_i} \end{bmatrix} \end{matrix} \tag{7}$$

where P_i denotes the number of results for discernment frame 1 and P denotes the number of results for discernment frame 2.

Based on the rule/utility information transform technology [28], the input information is transformed into the confidence distribution form under discernment framework 1, as shown below:

$$E^i(x_i^*) = \left\{(H_{k,i}, \eta_{k,i}), k = 1, 2, \ldots, P_i; (H_\Theta, \eta_{\Theta,i})\right\} \tag{8}$$

where x_i^* represents the ith input indicator. $H_{k,i}$ represents the kth referential grade of the ith indicator in discernment framework 1. $\eta_{k,i}$ denotes the belief degree assigned to any individual reference grade in the discernment framework 1, $0 \leq \eta_{n,i} \leq 1 (n = 1, 2, \ldots, P_i, i = 1, 2, \ldots, J)$. If the quantitative input information is x_i, then $\eta_{k,i}$ can be calculated as follows:

$$\begin{cases} \eta_{k,i} = \frac{H_{k+1,i} - x_i^*}{H_{k+1,i} - H_{k,i}} & H_{k,i} \leq x_i^* \leq H_{k+1,i} \\ \eta_{k+1,i} = 1 - \eta_{k,i} & H_{k,i} \leq x_i^* \leq H_{k+1,i} \\ \eta_{m,i} = 0 & m = 1, 2, \ldots, P_i, m \neq k, k+1 \end{cases} \tag{9}$$

where $H_{k+1,i}$ and $H_{k,i}$ represent the maximum and minimum referential values, respectively.

Finally, based on the transformation matrix $\mathbf{Z}_i$, the belief distribution of the input indicator x_i can be mapped from discernment framework 1 to discernment framework 2 as follows:

$$\widetilde{E}^i(x_i^*) = \left\{(F_{n,i}, \rho_{n,i}), n = 1, 2, \ldots, P; (F_\Theta, \rho_{\Theta,i})\right\} \tag{10}$$

where $0 \leq \rho_{n,i} \leq 1 (n = 1, 2, \ldots, P, i = 1, 2, \ldots, J)$ denotes the belief degree assigned to the nth result grade. $\rho_{\Theta,i} = 1 - \sum_{n=1}^{P} \rho_{n,i}$ denotes global ignorance. $\rho_{n,i}$ and $\rho_{\Theta,i}$ can be calculated as follows:

$$p_i = \mathbf{Z}_i \times n_i \tag{11}$$

$$\rho_{\Theta,i} = 1 - \sum_{n=1}^{P} \rho_{n,i} = \eta_{\Theta,i} \tag{12}$$

where $p_i = [\rho_{1,i}, \rho_{2,i}, \ldots, \rho_{P,i}]$ is the new belief degree after transformation, $n_i = [\eta_{1,i}, \eta_{2,i}, \ldots \eta_{P_i,i}]$ is the belief degree under discernment framework 1, and $\mathbf{Z}_i$ denotes the transform matrix for the ith indicator.

3.2. Reasoning Process of the E-BRB Model

Too many input attributes of the BRB model will lead to the problem of combinatorial explosion. The ER algorithm can analyze a large amount of uncertain information, which reduces the complexity of the emotional assessment indicators and obtains credible fusion results. Multiple emotional indicators are fused and input into BRB, which can effectively solve the combinatorial explosion problem.

Let us assume that input information x_i to the ER algorithm is quantitative information. The rule/utility-based transformation technique can equivalently transform the input information into the belief distribution shown in Equation (8). As described in Section 3.1, when the input reference grade does not match the result grade, the input information is transformed to the belief distribution under discernment framework 2 by the transform matrix $\mathbf{Z}_i$, as shown in Equation (10).

The evidence weight q_i is determined based on expert knowledge, which meets $0 \leq q_i \leq 1$. The fusion process using the ER algorithm can be described as follows:

$$\varphi_n = \frac{v \left[\prod_{k=1}^{J} (q_k \rho_{n,k} + 1 - q_k \sum_{j=1}^{P} \rho_{j,k}) - \prod_{k=1}^{J} (1 - q_k \sum_{j=1}^{P} \rho_{j,k}) \right]}{1 - v \left[\prod_{k=1}^{J} (1 - q_k) \right]} \tag{13}$$

$$v = \left[\sum_{n=1}^{P} \prod_{k=1}^{J} (q_k \rho_{n,k} + 1 - q_k \sum_{j=1}^{P} \rho_{j,k}) - (N-1) \prod_{k=1}^{J} (1 - q_k \sum_{j=1}^{P} \rho_{j,k}) \right]^{-1} \tag{14}$$

where φ_n denotes the belief degree of the nth result grade F_n after fusing the input indicators. $0 \leq \varphi_n \leq 1, \sum_{n=1}^{P} \varphi_n = 1$. Let us suppose that the utility of the assessment grade F_n is $u(F_n)$, and the expected utility is calculated as follows:

$$y_i = \sum_{n=1}^{P} u(F_n) \varphi_n \tag{15}$$

where y_i represents the fusion result of the ER algorithm.

Fusion through the ER algorithm can reduce the complexity of input emotional indicators $x_i, i = 1, \ldots, J$ and then use the fusion results as the input of the BRB model. After obtaining the fusion result $y_i, i = 1, \ldots, M$, the matching degree to the kth rule can be described by the following formula:

$$a_i^k = \begin{cases} \frac{A_i^{l+1} - y_i}{A_i^{l+1} - A_i^l} & k = l, A_i^l \leq y_i \leq A_i^{l+1} \\ \frac{y_i - A_i^l}{A_i^{l+1} - A_i^l} & k = l+1 \\ 0 & k = 1, 2, \ldots, L, k \neq l, l+1 \end{cases} \tag{16}$$

where a_i^k is the matching degree of the input information to the ith attribute in the kth rule. y_i denotes the input data for the ith antecedent attribute, which is the fusion result of the ER algorithm. A_i^l and A_i^{l+1} represent the referential values of the ith attribute in the two adjacent activation rules, the lth rule and the $(l+1)$th rule, respectively.

Then, the total matching degree, including matching degree a_i^k and attribute weight δ_i, can be calculated by

$$\bar{\delta}_i = \frac{\delta_i}{\max_{i=1,2,\ldots,M} \{\delta_i\}}, 0 \leq \bar{\delta}_i \leq 1 \tag{17}$$

$$a_k = \prod_{i=1}^{M} (a_i^k)^{\bar{\delta}_i} \tag{18}$$

where $\bar{\delta}_i$ denotes the weight of the ith attribute after normalization. M is the number of attributes. a_k is the total matching degree of the kth rule.

After obtaining the total matching degree, the activation weight of the kth rule is calculated. The calculation process is described by Equation (19):

$$\omega_k = \frac{\theta_k a_k}{\sum\limits_{l=1}^{L} \theta_l a_l}, k = 1, 2, \ldots, L \tag{19}$$

where θ_k denotes the weight of the kth rule. ω_k represents the activation weight of the kth rule.

When some rules are activated, the belief degree of y_i to different emotional grades can be calculated by the ER algorithm. The calculation process of the algorithm is shown in Equations (20) and (21):

$$\beta_n = \frac{\mu \left[\prod\limits_{k=1}^{L} \left(\omega_k \beta_{n,k} + 1 - \omega_k \sum\limits_{j=1}^{N} \beta_{j,k} \right) - \prod\limits_{k=1}^{L} \left(1 - \omega_k \sum\limits_{j=1}^{N} \beta_{j,k} \right) \right]}{1 - \mu \left[\prod\limits_{k=1}^{L} \left(1 - \omega_k \right) \right]} \tag{20}$$

$$\mu = \left[\sum\limits_{n=1}^{N} \prod\limits_{k=1}^{L} \left(\omega_k \beta_{n,k} + 1 - \omega_k \sum\limits_{j=1}^{N} \beta_{j,k} \right) - (N-1) \prod\limits_{k=1}^{L} \left(1 - \omega_k \sum\limits_{j=1}^{N} \beta_{j,k} \right) \right]^{-1} \tag{21}$$

where β_n denotes the belief degree of the nth emotional grade D_n, which satisfies $0 \leq \beta_n \leq 1$ and $\sum_{n=1}^{N} \beta_n = 1$.

The final belief degree generated after merging rules can be expressed as follows:

$$S(y_i) = \{(D_n, \beta_n); n = 1, 2, \ldots, N\} \tag{22}$$

where y_i denotes the input of the ith attribute. $S(\cdot)$ represents the nonlinear function modeled by E-BRB. The final output results are calculated according to the utility formula. $u(D_n)$ denotes the utility of D_n. The expected utility of y_i is described as

$$u(S(y)) = \sum\limits_{n=1}^{N} u(D_n) \beta_n \tag{23}$$

where $u(S(y_i))$ denotes the final results of the E-BRB model.

3.3. Optimization of the E-BRB Model

The parameters of the initial E-BRB model are determined by expert knowledge and may not be accurate due to the limitation of ambiguous knowledge representation. For more accurate parameters and results, we introduce an optimization model in this subsection to improve the accuracy of the model.

In the E-BRB model, the evidence weights, transformation matrix, attribute weights, rule weights, and belief degrees are the parameters that need to be optimized and should satisfy the following constraints.

- The evidence weights. The initial evidence weight q_i is determined by the expert and is subject to the constraints shown below:

$$0 \leq q_i \leq 1, i = 1, 2, \ldots, J \tag{24}$$

- The transform matrix. The initial value of the transformation matrix $\mathbf{Z}_i = \left[z_{j,k} \right]_{P \times P_i}$ of the ith indicator is given by the expert and must satisfy the following constraints:

$$0 \leq z_{j,k} \leq 1 \tag{25}$$

$$\sum_{j=1}^{P} z_{j,k} = 1 \tag{26}$$

- The attribute weights. Attribute weights can reflect the relative importance of attributes. The initial attribute weight δ_i is determined by experts, and the constraint conditions are as follows:

$$0 \leq \delta_i \leq 1, i = 1, 2, \ldots, M \tag{27}$$

- The rule weights. For the kth rule, its initial weight θ_k is determined by experts and is subject to the constraints shown below:

$$0 \leq \theta_k \leq 1, k = 1, 2, \ldots, L \tag{28}$$

- The belief degrees. In the kth rule, the belief degree $\beta_{n,k}$ corresponding to the result level D_n should satisfy the following constraint:

$$0 \leq \beta_{n,k} \leq 1, n = 1, 2, \ldots, N, k = 1, 2, \ldots, L \tag{29}$$

The sum of the belief degree in the results should satisfy the following formula. The equality sign holds if the kth rule is complete.

$$\sum_{n=1}^{N} \beta_{n,k} \leq 1, k = 1, 2, \ldots, L \tag{30}$$

Then, we utilize the mean square error (*MSE*) to measure the pros and cons of the E-BRB model, and its calculation equation can be expressed as:

$$MSE(q_i, \mathbf{Z}_i, \delta_i, \theta_k, \beta_{n,k}) = \frac{1}{T} \sum_{t=1}^{T} (\hat{u}(t) - u(t))^2 \tag{31}$$

where T is the number of model input data. $\hat{u}(t)$ represents the output value of the model. $u(t)$ denotes the actual output value.

Finally, the optimization objective function and constraints are as follows:

$$
\begin{aligned}
&\min MSE(\Omega) \\
&\quad s.t.\ 0 \leq q_i \leq 1, i = 1, 2, \ldots, J \\
&\qquad 0 \leq z_{j,k} \leq 1, j = 1, 2, \ldots, P, k = 1, 2, \ldots, P_i \\
&\qquad 0 \leq \delta_i \leq 1, i = 1, 2, \ldots, M \\
&\qquad 0 \leq \theta_i \leq 1, i = 1, 2, \ldots, L \\
&\qquad 0 \leq \beta_{n,k} \leq 1, n = 1, 2, \ldots, N; k = 1, 2, \ldots, L \\
&\qquad \sum_{j}^{P} z_{j,k} = 1 \\
&\qquad \sum_{n=1}^{N} \beta_{n,k} = 1
\end{aligned}
\tag{32}
$$

Formula (32) shows that the parameter optimization of the E-BRB model is a single objective multi-constraint optimization problem. In E-BRB, the constrained problem is a strongly constrained problem. Under the constraint condition, the feasible region of the solution is much smaller than the solution space. Given the superiority of P-CMA-ES in addressing high-dimensional non-linear optimization problems [29], it is utilized as the

optimization algorithm in this paper. The P-CMA-ES algorithm is developed from the CMA-ES algorithm [30,31]. The original algorithm finds the optimal solution by simulating biological evolution. The P-CMA-ES algorithm adds a projection operation after the selection operation of the original algorithm to map the solutions that do not meet the constraints back to the feasible region [32]. The optimization process of the P-CMA-ES algorithm is shown in Figure 3.
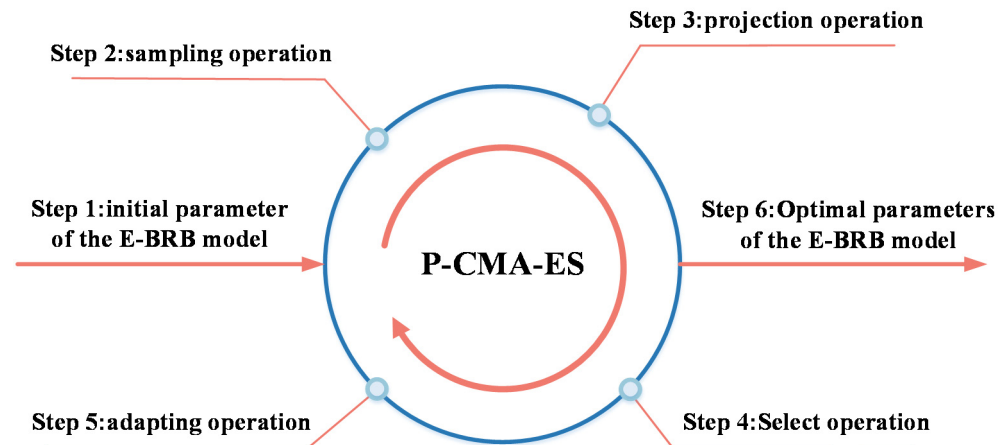

Figure 3. Parameter optimization process of the P-CMA-ES algorithm.

As shown in Figure 3, The P-CMA-ES optimization process can be divided into six steps. The specific details are described as follows:

Step 1: Give the initial parameters $w^0 = \Omega^0$. Ω^0 denotes the initial parameter vector to be optimized in the EBRB model. $\Omega = \{q_1, \ldots, q_J, \mathbf{Z}_1, \ldots \mathbf{Z}_J, \delta_1, \ldots, \delta_M, \theta_1, \ldots, \theta_L, \beta_{1,1}, \ldots, \beta_{N,L}\}$. Determine the initial parameters of the P-CMA-ES algorithm, including population size λ, and offspring population size τ.

Step 2: The sampling operation is performed, and the initial population is generated based on the normal distribution with the initial solution as the expected value. The specific process can be described as follows:

$$\Omega_i^{g+1} \sim w^g + \varepsilon^g \mathbb{N}(0, C^g) \tag{33}$$

where the $i(i = 1, \ldots, \lambda)$th solution in the $(g + 1)$ generation is represented as Ω_i^{g+1}. w^g represents the mean of the offspring population in the gth deneration. ε represents the evolutionary steps. $\mathbb{N}(\cdot)$ represents the normal distribution. The covariance matrix of the gth generation population is represented as C^g.

Step 3: The projection operation is executed on the solution that does not satisfy the constraint. The solution is projected into a hyperplane, which is the feasible region of the equality constraint. According to formula (32), there are $N + 1$ equality constraints in the E-BRB model, and each equality constraint contains N variables. The hyperplane can be denoted as $R_e \Omega_i^g (1 + v_e \times (m - 1) : v_e \times m) = 1$. $v_e = (1, \ldots, N)$ and $m = 1, \ldots, N + 1$ represent the number of variables that are constrained by the equation constraint and the number of equality constraints in solution Ω_i^g, respectively. $R_e = [1 \ldots 1]_{1 \times N}$ represents the parameter vector of the equation. The projection operation can be described as follows:

$$\begin{aligned} \Omega_i^{g+1}(1 + v_e \times (m - 1) : v_e \times m) \quad &= \Omega_i^{g+1}(1 + v_e \times (m - 1) : v_e \times m) - R_e^T \times (R_e \times R_e^T)^{-1} \\ &\times \Omega_i^{g+1}(1 + v_e \times (m - 1) : v_e \times m) \times R_e \end{aligned} \tag{34}$$

The solution processed by the projection operation may exceed the boundary constraint of the solution space. To solve this problem, the extra values of the equality constraint variables should be equally assigned to other variables.

Step 4: Perform selection and recombination operations. Select τ optimal solution according to the fitness function. Update the mean by Equation (35):

$$w^{g+1} = \sum_{i=1}^{\tau} h_i \Omega_{i:\lambda}^{g+1} \tag{35}$$

where h_i denotes the weight coefficient of the ith solution.

Step 5: Perform adaptive operations to update the covariance matrix to obtain the range and direction of the population search. The calculation process is shown in the following Equations:

$$C^{g+1} = (1 - a_1 - a_2)C^g + a_1 p_c^{g+1} \left(p_c^{g+1} \right)^T + a_2 \sum_{i=1}^{\tau} h_i \left(\frac{\left(\Omega_{i:\lambda}^{g+1} - w^g \right)}{\varepsilon^g} \right) \left(\frac{\left(\Omega_{i:\lambda}^{g+1} - w^g \right)}{\varepsilon^g} \right)^T \tag{36}$$

$$p_c^{g+1} = (1 - a_c)p_c^g + \sqrt{a_c(2 - a_c) \left(\sum_{i=1}^{\tau} h_i^2 \right)^{-1}} \frac{w^{g+1} - w^g}{\varepsilon^g} \tag{37}$$

$$\varepsilon^{g+1} = \varepsilon^g \exp\left(\frac{a_\sigma}{d_\sigma} \left(\frac{\|p_\sigma^{g+1}\|}{E\|\mathbb{N}(1, I)\|} - 1 \right) \right) \tag{38}$$

$$p_\sigma^{g+1} = (1 - a_c)p_\sigma^g + \sqrt{a_\sigma(2 - a_\sigma) \left(\sum_{i=1}^{\tau} h_i^2 \right)^{-1}} C^{(g) - \frac{1}{2}} \frac{w^{g+1} - w^g}{\varepsilon^g} \tag{39}$$

where a_1 and a_2 represent the learning rate. a_σ denotes the backward time horizon. p_c^{g+1} represents the evolution path of the covariance matrix in the $g + 1$th generation. d_σ is the damping coefficient. $E\|\mathbb{N}(1, I)\|$ is the expectation of the normal distribution $\mathbb{N}(1, I)$. p_σ^{g+1} represents the conjugate evolution step in the $g + 1$th generation.

Step 6: Repeat steps 2 to 5 until the best solution $\Omega_{optimal}$ is found.

3.4. Modeling Method of Learning Emotion Assessment Based on E-BRB

The modeling method of learning emotion assessment based on E-BRB is introduced in this subsection. Based on the above analysis, the implementation of the model mainly includes three parts: model construction, parameter training, and model testing. The details are summarized as follows.

First, the initial E-BRB model is constructed based on the sample data and the initial parameters given by experts.

The second is the training part. Considering the influence of the limited expert knowledge on the model's accuracy, the parameters given in Section 3.3 are trained by the optimization model in this part. The training data are used as input to the E-BRB model, and the optimized E-BRB model is obtained after this part.

Finally, there is the testing part. After the training part, we obtain the optimal parameters of the model, including the transformation matrix $\mathbf{Z}_i$, evidence weights q_i, attribute weights δ_i, rule weights θ_k, and belief degrees $\beta_{n,k}$. The estimated output is obtained by the E-BRB model using the testing data as input.

Based on the above discussion, the implementation of the E-BRB model is shown in Figure 4, which can be summarized as follows:
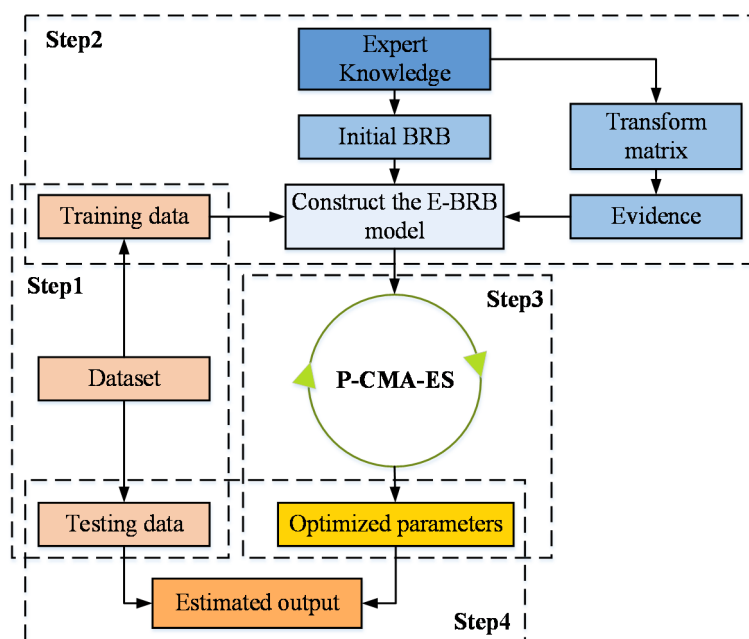
Figure 4. The implementation process of the E-BRB model.

Step 1: Collect and divide data into training data and testing data. The division method can be the random split or other methods.

Step 2: Build an initial E-BRB model based on expert knowledge.

Step 3: After obtaining the training data and the initial values Ω of the E-BRB model, the E-BRB model can be trained in the training part. The P-CMA-ES algorithm is used to obtain the optimized model according to the optimization objectives. The optimization steps are performed recursively until the optimal solution $\Omega_{optimal}$ is obtained.

Step 4: The testing data are tested on the optimized E-BRB model to obtain the final output of the model. The accuracy of the model is represented by the *MSE* value.

Step 4.1: The transformed belief distribution is obtained by using Equations (11) and (12) and fused by Equations (13)–(15).

Step 4.2: The matching degree and the activation weight are obtained according to Equations (16)–(19).

Step 4.3: The ER algorithm is utilized to aggregate the activated belief rules. Calculate the final output of the E-BRB model using Equation (23).

Step 4.4: The *MSE* value is calculated by Equation (31), which reflects the modeling accuracy of the E-BRB model.

4. Case Study

The scientific learning emotions of learners may have a negative impact on their scientific learning performance. It is necessary to evaluate learners' scientific learning emotions to explain the mechanism of learning emotions. A case of student scientific emotion assessment is presented in this section to verify the effectiveness of the proposed model. This section is divided into the following four parts. In Section 4.1, the basic definition of the experiment in this case study is introduced. In Section 4.2, a scientific emotion assessment model is constructed. In Section 4.3, the training and testing of the model are presented. In Section 4.4, comparative experiments are conducted. The experimental analysis is discussed in Section 4.5.

4.1. The Basic Definition of the Experiment

Data for this case study come from the context questionnaire scale of the Iranian region that participated in the eighth grade of TIMSS2019 [33]. The TIMSS2019 dataset collects and summarizes data in the Likert scale format, with a total of 17 indicators. The number

and content of items are listed in Tables 1 and 2. The Likert scale requires respondents to indicate their degree of agreement with a declarative statement. However, the Likert data are somewhat ambiguous in terms of data quality and potential variable assessment. For example, such data may collect incomplete information when a particular problem does not apply to respondents. In this experiment, 400 sets of samples are selected, of which 280 sets of samples are used for training parameters, and 120 sets of samples are used for model testing.

Table 1. The number of items.

Name	Number of Items
Confidence in science	8
Value science	9

Table 2. Details of the dataset.

Number	Items
1	I usually do well in science
2	Science is more difficult for me than for many of my classmates
3	Science is not my strengths
4	I learn things quickly in science
5	I am good at working out difficult science problems
6	My teacher tells me I am good at science
7	Science is harder for me than any other subject
8	Science makes me confused
9	I think learning science will help me in my daily life
10	I need science to learn other school subjects
11	I need to do well in science to get into the university of my choice
12	I need to do well in science to get the job I want
13	I would like a job that involves using science
14	It is important to learn about science to get ahead in the world
15	Learning science will give me more job opportunities when I am an adult
16	My parents think that it is important that I do well in science
17	It is important to do well in science

4.2. Construction of the E-BRB Model

Two key properties were identified through the analysis of the dataset. They are the degree of self-confidence and the degree of identification, respectively. Scientific self-confidence reflects the degree to which individuals think they are capable of scientific disciplines, and the degree of identification reflects the degree to which individuals attach importance to scientific disciplines. The indicators in the dataset are divided into the degree of confidence (ϖ_1) and the degree of identification (ϖ_2), which are x_1–x_8 and x_9–x_{17} in the dataset, respectively. The data in the dataset are summarized and collected in a four-point scale format, where 1 = strongly agree, 2 = somewhat agree, 3 = somewhat disagree, and 4 = strongly disagree. According to the actual situation, the result grade of ϖ_1 can be divided into $F^1 = \{F_1, F_2, F_3, F_4, F_5\}$ = {unconfident (U), less confident (LC), a little confident (LC), quite confident (QC), very confident (VC)}. The result grade of ϖ_2 can be divided into $F^2 = \{F_1, F_2, F_3, F_4\}$ = {unimportant (U), less important (LI), slightly important (SI), very important (VI)}. However, according to the statements in the scale items, the reference grades for the input indicators ϖ_1 can be divided into $H^1 = \{H_1, H_2, H_3, H_4\}$ = {very anxious (VA), slightly anxious (SA), less anxious (LA), not anxious (NA)}. The reference grades for the input indicators ϖ_2 can be divided into $H^2 = \{H_1, H_2, H_3, H_4, H_5\}$ = {weak (W), little weak (LW), middle (M), little strong (LS), strong (S)}.

After determining the antecedents and outcome parameters of the rules, the transformation matrix can be established in Tables 3 and 4 through Formula (6). The sum of the belief degree of the results in the transformation matrix is 1.

Table 3. The parameters of the transformation matrix A_1.

H^1	$\{F_1, F_2, F_3, F_4, F_5\}$
H_1	(0.9, 0.1, 0, 0, 0)
H_2	(0.1, 0.7, 0.2, 0,0)
H_3	(0, 0.05, 0.35, 0.6, 0)
H_4	(0, 0, 0, 0.05, 0.95)

Table 4. The parameters of the transformation matrix A_2.

H^2	$\{F_1, F_2, F_3, F_4\}$
H_1	(1, 0, 0, 0)
H_2	(0.6, 0.4, 0, 0)
H_3	(0.2, 0.7, 0.1, 0)
H_4	(0, 0.15, 0.85, 0)

On the basis of Tables 3 and 4, the transformation matrixes A_1 and A_2 can be described as follows:

$$
A_1 = \begin{bmatrix} 0.9 & 0.1 & 0 & 0 \\ 0.1 & 0.7 & 0.05 & 0 \\ 0 & 0.2 & 0.35 & 0 \\ 0 & 0 & 0.6 & 0.05 \\ 0 & 0 & 0 & 0.95 \end{bmatrix} A_2 = \begin{bmatrix} 1 & 0.6 & 0.2 & 0 & 0 \\ 0 & 0.4 & 0.7 & 0.15 & 0 \\ 0 & 0 & 0.1 & 0.85 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{40}
$$

According to Equations (8)–(12), the input information can be transformed into a belief distribution. For instance, let us suppose the value of the indicator x_1 is 2, and the belief distribution of formula (10) can be expressed as $\widetilde{S}(x^*) = \{(F_1, 0), (F_2, 0.0375), (F_3, 0.2625), (F_4, 0.4625), (F_5, 0.2375)\}$.

After obtaining the belief distribution, the ER algorithm is used for evidence fusion. Since the data used in the experiment are Likert scale data, the same initial weight is given to all indicators, namely, $q_i = 0.9$. y_1 denotes the result of the ϖ_1 attribute after ER algorithm fusion and y_2 denotes the result of the ϖ_2 attribute after ER algorithm fusion. The referential points and referential values for y_1 and y_2 are given in Tables 5 and 6 in combination with the results obtained. In this paper, we use five points for y_1: very small (VS), small (S), middle (M), large (L), and very large (VL). Similarly, we use four points for y_2: very small (VS), small (S), middle (M), and large (L). For the consequent attribute, emotion state, four referential points are used: strong negative (SN), weak negative (WN), weak positive (WP), and strong positive (SP), as shown in Table 7. y_1 has five reference points and y_2 has four reference points. According to the Cartesian product, there are 20 rules in the model. The initial parameters of the model are determined by experts, which are given in Table 8.

Table 5. Referential points and values for y_1.

Referential Point	Referential Value
VS	3
S	6
M	9
L	12
VL	15

Table 6. Referential points and values for y_2.

Referential Point	Referential Value
VS	4
S	6.5
M	10
L	13

Table 7. Referential points and values for the emotional state.

Referential Point	Referential Value
SN	3
WN	8
WP	10
SP	14

Table 8. Initial belief rules.

Rule Number	Rule Weight	y_1 and y_2	Attitude Distribution $\{D_1, D_2, D_3, D_4\}$ = {3, 8, 10, 14}
1	1	VS AND VS	{0, 0, 0, 0}
2	1	VS AND S	{0.7, 0.3, 0, 0}
3	1	VS AND M	{0, 0.8, 0.2, 0}
4	1	VS AND L	{0, 0.45, 0.55, 0}
5	1	S AND VS	{0.35, 0.65, 0, 0}
6	1	S AND S	{0.1, 0.9, 0, 0}
7	1	S AND M	{0, 0.65, 0.35, 0}
8	1	S AND L	{0, 0.3, 0.7, 0}
9	1	M AND VS	{0.2, 0.8, 0, 0}
10	1	M AND S	{0, 0.85, 0.15, 0}
11	1	M AND M	{0, 0.4, 0.6, 0}
12	1	M AND L	{0, 0.15, 0.85, 0}
13	1	L AND VS	{0, 0.7, 0.3, 0}
14	1	L AND S	{0, 0.45, 0.55, 0}
15	1	L AND M	{0, 0.1, 0.7, 0.2}
16	1	L AND L	{0, 0, 0.65, 0.35}
17	1	VL AND VS	{0, 0, 0.55, 0.45}
18	1	VL AND S	{0, 0, 0.3, 0.7}
19	1	VL AND M	{0, 0, 0.1, 0.9}
20	1	VL AND L	{0, 0, 0, 1}

4.3. Training and Testing for the E-BRB Model

After the construction of the model, to reduce the uncertainty caused by expert knowledge, the parameters need to be optimized. In this section, the E-BRB model is trained based on the acquired data. A total of 159 parameters are trained in the model training part, including the transformation matrix, evidence weights, attribute weights, rule weights, and rule output belief degrees. There are 400 groups of experimental data in this paper, which belong to small-scale datasets. Based on the common proportion of small-scale datasets, 400 groups of data are randomly divided according to the ratio of the training set to the testing set 7:3, of which 280 groups are used for training data, and the remaining 120 groups are used as testing data. The number of iterations in the P-CMA-ES algorithm is 25 and 400, respectively.

The optimization weights of attribute one and attribute two are 0.7178 and 0.8148, respectively. The optimized weights of the evidence and the E-BRB model are presented in Tables 9–11. The optimized transformation matrixes are as follows:

$$A_1 = \begin{bmatrix} 0.3558 & 0.0217 & 0.1061 & 0.0017 \\ 0.1341 & 0.3005 & 0.1720 & 0.1698 \\ 0.2248 & 0.2844 & 0.1846 & 0.0811 \\ 0.154 & 0.3667 & 0.4972 & 0.475 \\ 0.1313 & 0.0268 & 0.0402 & 0.2724 \end{bmatrix} \tag{41}$$

$$A_2 = \begin{bmatrix} 0.1918 & 0.3361 & 0.2531 & 0.4537 & 0.0894 \\ 0.1696 & 0.4012 & 0.1653 & 0.1884 & 0.6659 \\ 0.2061 & 0.1946 & 0.1085 & 0.3411 & 0.0393 \\ 0.4325 & 0.0680 & 0.4732 & 0.0168 & 0.2054 \end{bmatrix} \tag{42}$$

Table 9. Optimized weights of the relevant indicators in ϖ_1.

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
0.555	0.2618	0.5268	0.9428	0.5926	0.0572	0.4097	0.6092

Table 10. Optimized weights of the relevant indicators in ϖ_2.

x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
0.8179	0.3491	0.3374	0.2871	0.4918	0.6512	0.0709	0.5135	0.2192

Table 11. Optimized belief rules.

Rule Number	Rule Weight	y_1 and y_2	Attitude Distribution $\{D_1, D_2, D_3, D_4\}$ = {3, 8, 10, 14}
1	0.7894	VS AND VS	{0.1235, 0.2923, 0.2684, 0.3157}
2	0.9858	VS AND S	{0.1382, 0.0664, 0.0916, 0.7038}
3	0.0438	VS AND M	{0.1917, 0.721, 0.0233, 0.064}
4	0.7306	VS AND L	{0.3512, 0.4116, 0.1255, 0.1117}
5	0.8189	S AND VS	{0.1292, 0.4537, 0.2154, 0.2017}
6	0.2281	S AND S	{0.044, 0.4131, 0.3107, 0.2323}
7	0.7077	S AND M	{0.7111, 0.2063, 0.0776, 0.0049}
8	0.7062	S AND L	{0.0622, 0.3248, 0.5661, 0.0469}
9	0.7033	M AND VS	{0.2933, 0.0654, 0.3704, 0.2709}
10	0.7589	M AND S	{0.0132, 0.7044, 0.2277, 0.0547}
11	0.7910	M AND M	{0.3945, 0.1609, 0.4333, 0.0112}
12	0.7219	M AND L	{0.1641, 0.5493, 0.141, 0.1457}
13	0.7971	L AND VS	{0.1105, 0.2097, 0.2793, 0.4005}
14	0.3508	L AND S	{0.0029, 0.0153, 0.0029, 0.9789}
15	0.6414	L AND M	{0.4455, 0.0188, 0.1099, 0.4259}
16	0.8209	L AND L	{0.0976, 0.5039, 0.0341, 0.3644}
17	0.4206	VL AND VS	{0.0529, 0.2121, 0.4366, 0.2984}
18	0.4715	VL AND S	{0.2393, 0.4456, 0.248, 0.0671}
19	0.3714	VL AND M	{0.4374, 0.0976, 0.2052, 0.2598}
20	0.3343	VL AND L	{0.0664, 0.2238, 0.2385, 0.4713}

To evaluate the performance of the model, mean square error (*MSE*), root mean square error (*RMSE*), and mean absolute error (*MAE*) are introduced to measure the performance of the model. These three metrics are the most commonly utilized performance evaluation criteria and have been used in this study. The formulas are as follows, where T is the number of model input data, $\hat{u}(t)$ represents the output value of the model, and $u(t)$ represents the true value.

- *MSE*

The calculation equation of *MSE* is shown in Equation (31). *MSE* is a more convenient way to measure the mean error. The smaller the *MSE* value, the better the accuracy of the model.

- *RMSE*

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{u}(t) - u(t))^2} \tag{43}$$

The *RMSE* value represents the standard deviation of the residual between the measured true value and the predicted value, which is the square root of the *MSE* value. It is more sensitive to outliers in data than *MAE*.

- *MAE*

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|\hat{u}(t) - u(t)| \tag{44}$$

The *MAE* value is the mean of the absolute error between the true value and the predicted value. In contrast, it is less sensitive to extreme values and has better robustness to outliers.

The comparison between the testing results of the learning emotion assessment model and the actual results is shown in Figure 5, where the true value is the learner's true emotional state score, and the predicted value is the output of the E-BRB model. The *MSE* value of the model output is 0.7963, the *RMSE* value is 0.8923, and the *MAE* value is 0.6729. As shown in Figure 5, the emotional state score estimated by the optimized E-BRB model fits well with the actual score. The E-BRB model optimized based on P-CMA-ES can accurately predict the emotional states of learners.
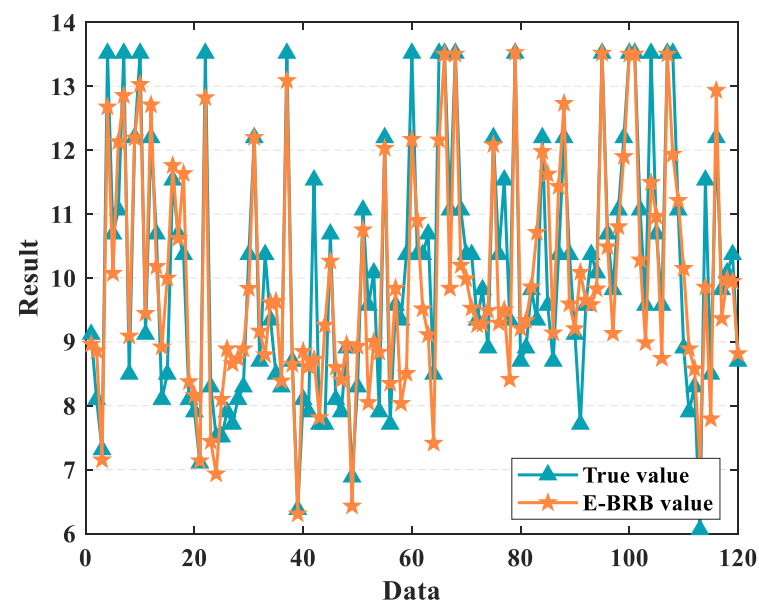
Figure 5. Fitting diagram of the E-BRB model.

4.4. Comparative Study

To demonstrate the effectiveness of the E-BRB model, we compare the proposed model with the backpropagation neural network (BPNN), K-nearest neighbor (KNN), SVM, extreme learning machine (ELM), random forest (RF), and decision tree (DT) models in this subsection. The number of training and testing is the same. BPNN and ELM are methods based on quantitative information. KNN uses proximity to classify or predict the grouping of individual data points. SVM attempts to find a hyperplane to segment samples. For RF, it is presumed that there are multiple trees, and each number represents an output. DT is a tree structure that can be a binary or non-binary tree. In the current

study, the above methods are commonly used assessment methods. The model experiment is implemented in Python and Matlab. The output results of the comparison model are shown in Figures 6–11. The *MSE* values of the six models are 0.8742, 0.959, 0.834, 0.9612, 0.882, and 1.007, respectively.
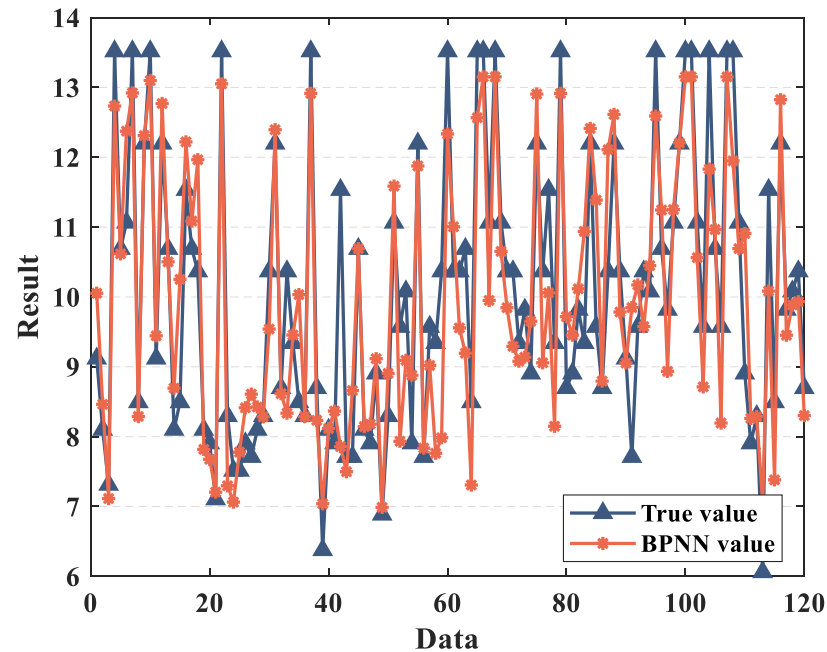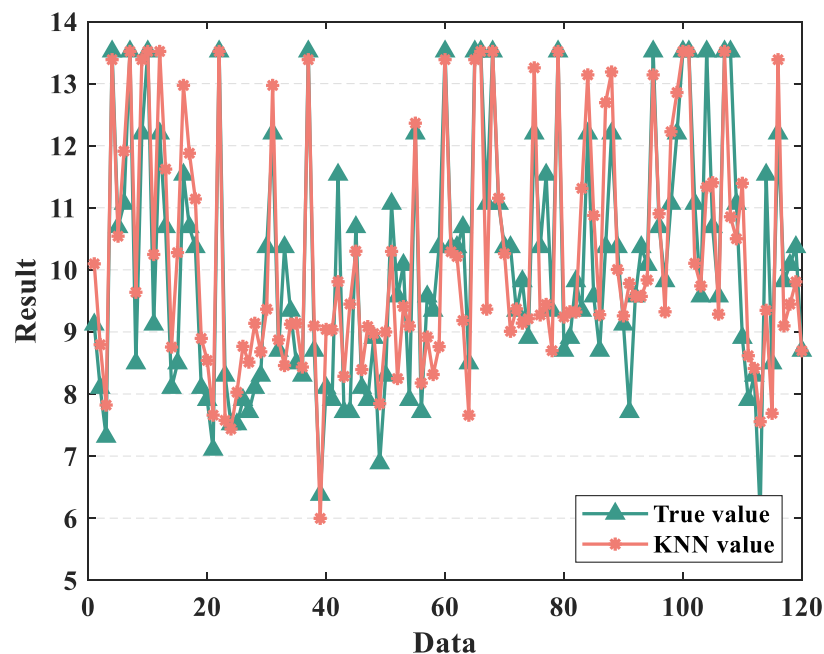

Figure 6. Fitting diagram of the BPNN model.

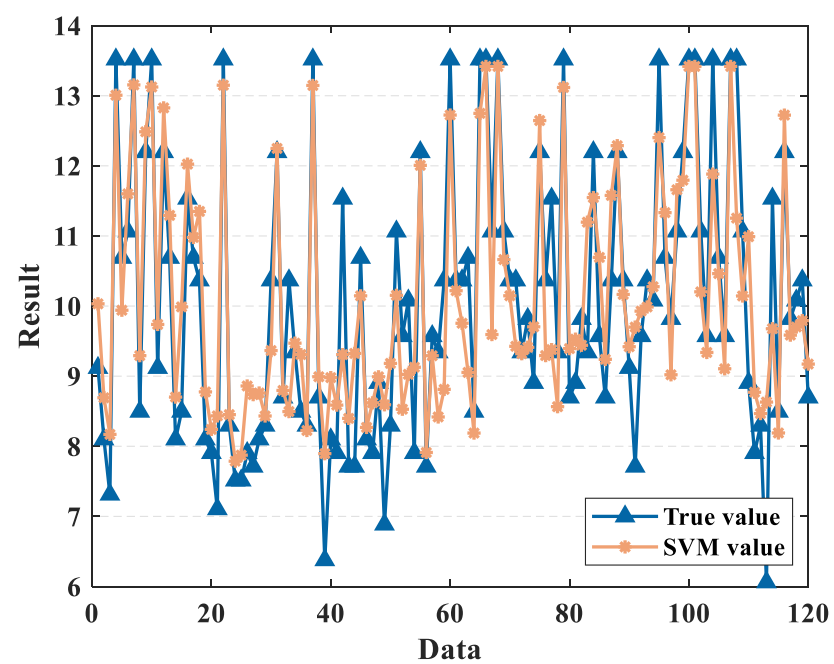**Figure 7.** Fitting diagram of the KNN model.
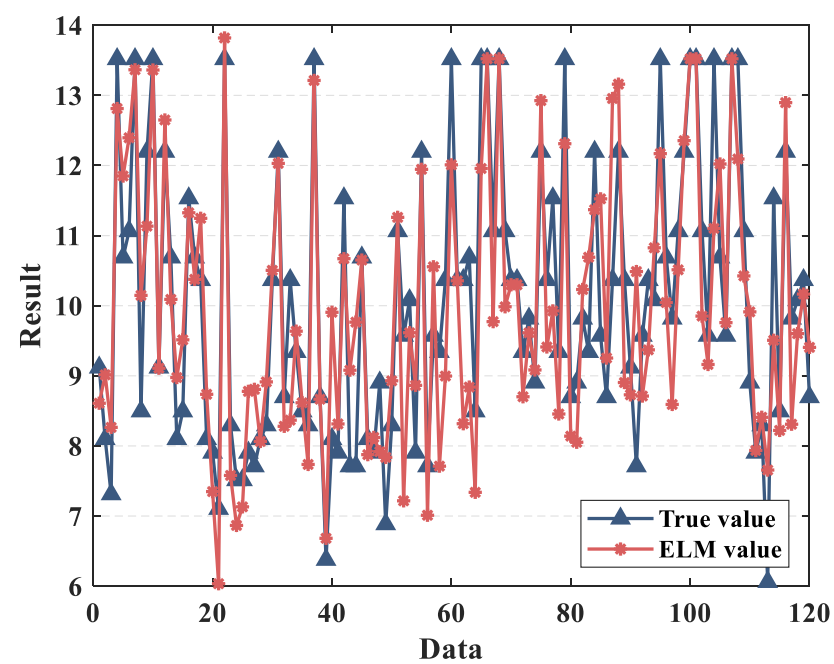
Figure 8. Fitting diagram of the SVM model.

Figure 9. Fitting diagram of the ELM model.
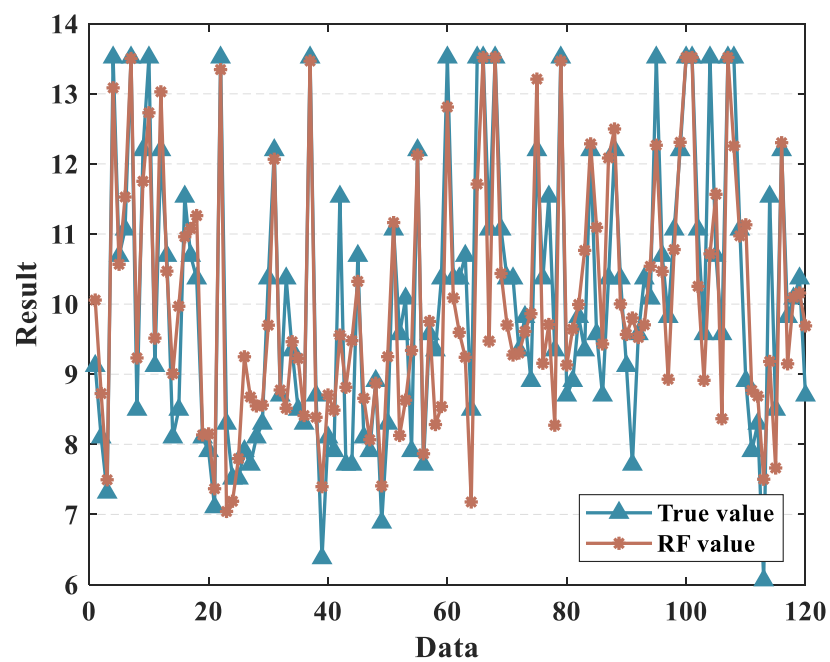
Figure 10. Fitting diagram of the RF model.
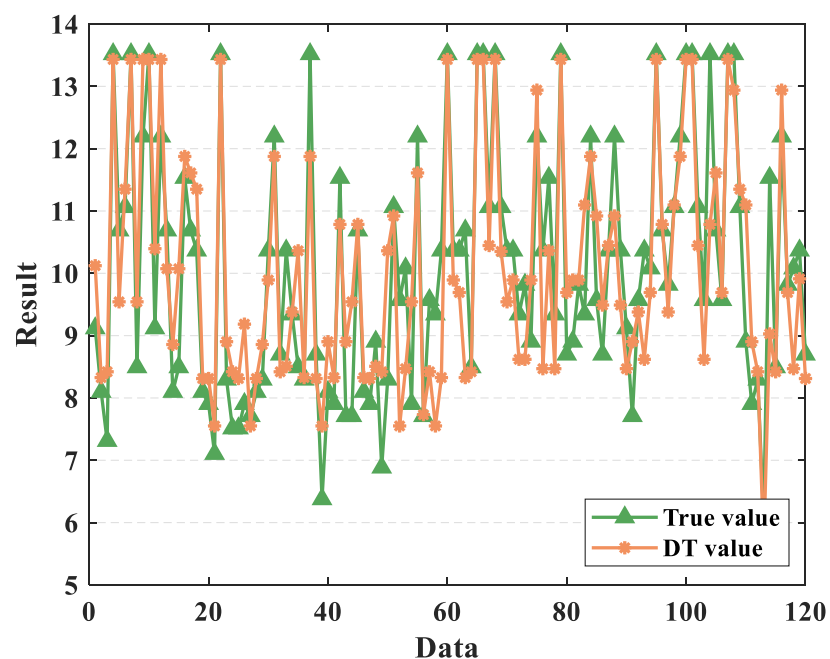
Figure 11. Fitting diagram of the DT model.

To demonstrate the robustness of the E-BRB model, we repeated the experiment 20 times with the same training and testing parts. The hyperparameters of the comparison model are given in Table 12. The average results for all methods are shown in Table 13. Figures 12–14 are the *MSE* values, *RMSE* values, and *MAE* values of repeated model experiments, respectively. The average *MSE, RMSE*, and *MAE* values of E-BEB are 0.8043, 0.8967, and 0.6801, respectively. It can be seen that the E-BRB model is more effective and robust than the other models in learning emotion assessment.

Table 12. Model hyperparameter settings.

Model	Parameters Settings
BPNN	hidden_layer_sizes = 5, learning_rate_init = 0.01, max_iter = 200
KNN	n_neighbors = 30, algorithm = 'auto', weights = 'uniform', leaf_size = 30, p = 2, metric = 'minkowski', metric_params = None, n_jobs = 1
SVM	kernel = 'rbf', C = 0.85
ELM	activation function = 'sig moidal', number of hidden neurons = 100
RF	n_estimators = 90, oob_score = True, random_state = 10
DT	splitter = 'best', min_samples_leaf = 10

Table 13. Comparison of *MSE*, *RMSE*, and *MAE* values.

Model	*MSE*	*RMSE*	*MAE*
E-BRB	0.8043	0.8967	0.6801
BPNN	0.9088	0.9529	0.7022
KNN	1.0142	1.006	0.737
SVM	0.9013	0.9491	0.699
ELM	1.066	1.0321	0.7781
RF	0.90007	0.9484	0.6971
DT	1.0702	1.0341	0.7738
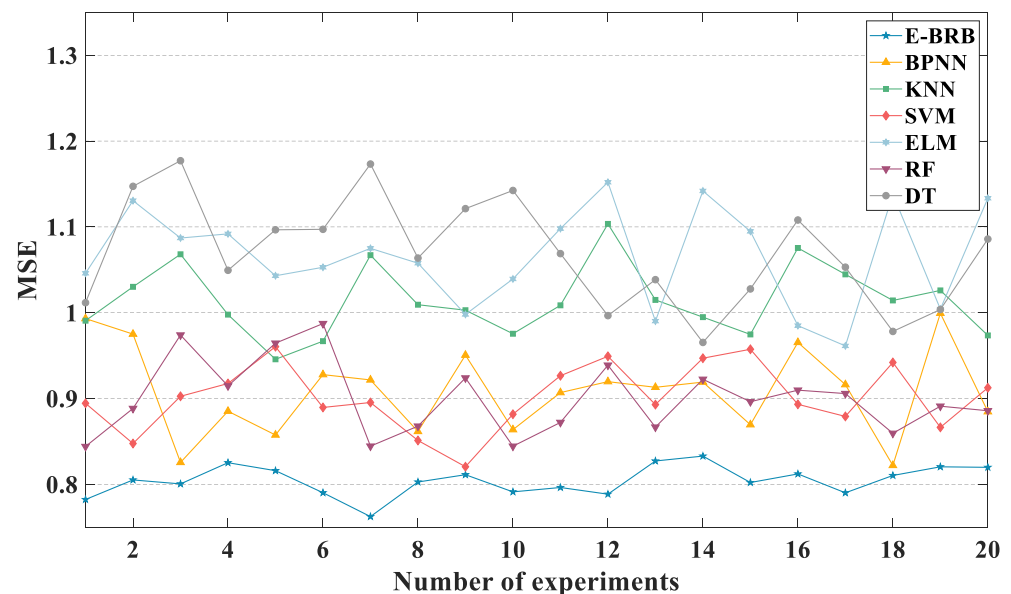
Figure 12. *MSE* comparison of the E-BRB, BPNN, KNN, SVM, ELM, RF, and DT models.

4.5. Discussion

According to Figures 10–12, the indicators of the trained E-BRB model are higher than those of several other models. Compared with other models, the *MSE* of E-BRBB improved by 11.49%, 20.69%, 10.76%, 24.54%, 10.64%, and 24.84%, respectively. Furthermore, the *RMSE* improved by 5.89%, 10.86%, 5.52%, 13.11%, 5.45%, and 13.28%, respectively. SVM and RF are pretty competitive with E-BRB in terms of *MAE* value. The average *MAE* of E-BRB is 0.6801. The *MAE* of SVM and RF are 0.699 and 0.6971, respectively.
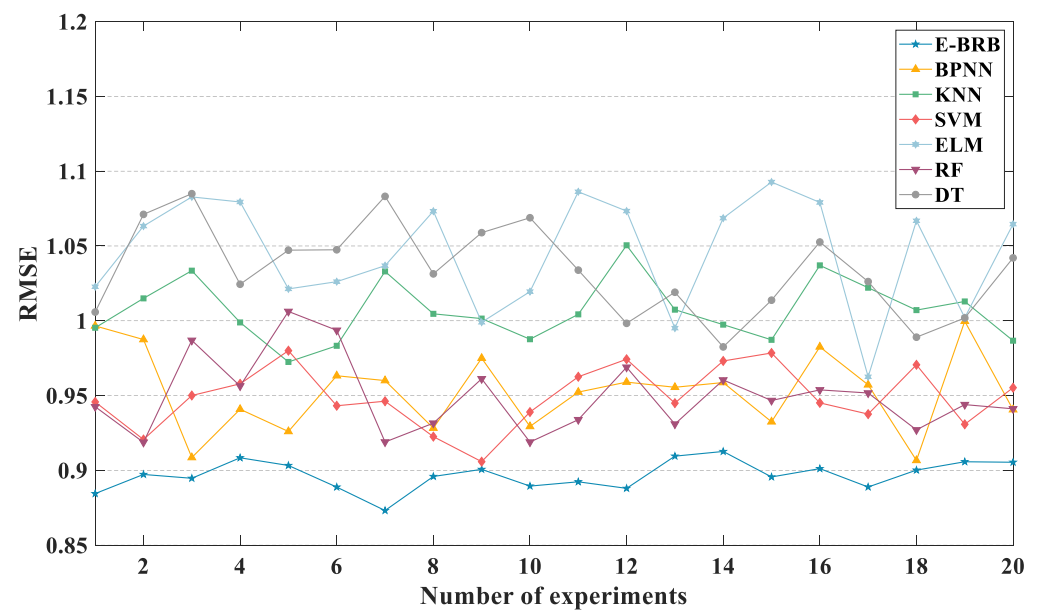
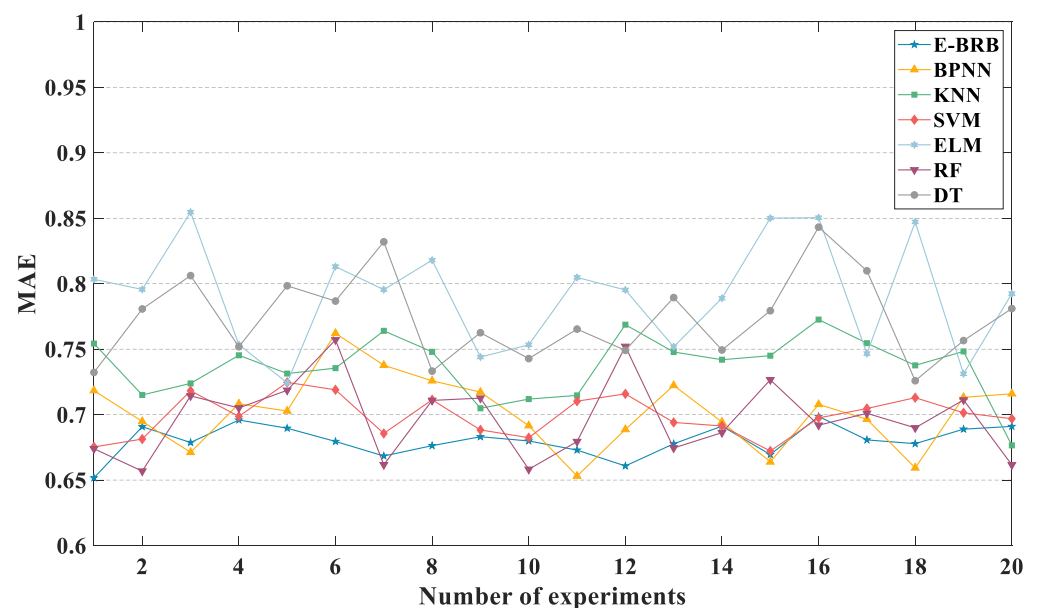**Figure 13.** *RMSE* comparison of the E-BRB, BPNN, KNN, SVM, ELM, RF, and DT models.

Figure 14. *MAE* comparison of the E-BRB, BPNN, KNN, SVM, ELM, RF, and DT models.

BPNN, SVM, and RF are the three most commonly used tools in learning emotion assessment, which are data-driven models. There are strengths to data-driven approaches in model derivation because they do not need to know the specific relationship between input and specific output results in advance. Nevertheless, the performance of the model varies considerably across different training rounds, even if the same dataset is used. This is explained by the fact that the performance of models that rely too much on data is determined by the training set. From the experimental results, BPNN, SVM, and RF have performed well. However, they cannot provide good interpretability. Due to the fact that they are black-box methods, the derivation process cannot be known with certainty. In contrast, the E-BRB model considers both expert knowledge and historical data. E-BRB utilizes expert knowledge to construct the initial model and employs historical data and optimization techniques to improve the accuracy of the model. The method allows for a greater expression of the relationship between input and specific output results. E-BRB presents

both the initial and optimization models with clear reasoning and optimization processes and greater transparency. Although DT has certain interpretability, its performance is not as good as E-BRB.

Through the analysis of the above experimental results, the following conclusions can be drawn:

- The parameters of the E-BRB model can be trained and optimized by the optimization algorithm, and the accuracy of the optimized model is better than other methods. From the average results of 20 repeated experiments, it can be seen that the E-BRB model has good robustness and better accuracy.
- The reasoning process based on the E-BRB model is traceable and can clearly explain the causal relationship between emotional indicators and emotional states. Therefore, the learning emotion assessment method based on E-BRB has better interpretability and credibility than other data-driven methods.

5. Conclusions

Aiming at the problem that the current learning emotion assessment model cannot take both model accuracy and model interpretability into account, a learning emotion assessment model based on E-BRB is proposed. Through the analyzing and processing of the emotional data generated by learners, their learning emotional state can be understood. When it is challenging to carry out learning intervention based on academic performance or learning behavior, the E-BRB emotional assessment model can help teachers carry out learning intervention from an emotional perspective and explore the mechanism of learning emotion in the learning process. The E-BRB model has two characteristics: (1) stronger ability of inference. (2) Better interpretability. Experimental results show that the E-BRB model has better performance in accuracy and stability. The inference process of E-BRB is visual, and the reasoning results are traceable. The model can be used for emotional assessment in the classroom environment, which is beneficial for teachers to master students' learning emotions and facilitate teaching. However, there are some limitations in this paper. There is insufficient consideration for the interpretability of the optimization model. When the E-BRB model is optimized, its interpretability may be damaged to some extent. For example, the optimized belief distribution is inconsistent with the actual emotional state, and the range of belief degrees is unreasonable.

Future research work can be carried out from the following two aspects: (1) Further improve the interpretability of the model. To ensure the interpretability of the model in the optimization process, how to make full use of expert knowledge and setting reasonable interpretability constraint criteria for the model optimization process of E-BRB needs further discussion. (2) Comprehensively measure learning emotions and promote learning intervention. Establish the relationship between learning emotion and individual characteristics such as cognitive ability, learning attitude, and learning behavior. On the other hand, multimodal data provide a more profound portrait of learners' relevant learning behaviors than a single data source. How to use multimodal data to build a complete data chain for accurate assessment and tracking feedback is the next step.

References

1. Zhou, J.; Ye, J.M.; Li, C. Multimodal Learning Affective Computing: Motivations, Frameworks and Suggestions. *E-Educ. Res.* **2021**, *42*, 26–32+46.
2. Wang, Y.Y.; Liu, S.Y.; Zheng, Y.H. Research on Emotional Perception of Learners in the Age of Intelligence: Connotation, Current Situation and Trend. *J. Distance Educ.* **2021**, *39*, 34–43.
3. Jin, X.Q.; Wang, L.L.; Yang, X.M. Construction of Online Learning Emotional Measurement Model based on Big Data. *Mod. Educ. Technol.* **2016**, *26*, 5–11.
4. Barrón Estrada, M.L.; Zatarain Cabada, R.; Oramas Bustillos, R.; Graff, M. Opinion mining and emotion recognition applied to learning environments. *Expert Syst. Appl.* **2022**, *150*, 113265. [CrossRef]
5. Ashwin, T.S.; Guddeti, R.M.R. Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks. *Educ. Inf. Technol.* **2020**, *25*, 1387–1415.
6. Bota, P.; Wang, C.; Fred, A.; Silva, H. Emotion Assessment Using Feature Fusion and Decision Fusion Classification Based on Physiological Data: Are We There Yet? *Sensors* **2020**, *20*, 4723. [CrossRef]
7. Chan, M.C.E.; Ochoa, X.; Clarke, D. Multimodal Learning Analytics in a Laboratory Classroom. In *Machine Learning Paradigms*; Virvou, M., Alepis, E., Eds.; Springer: Cham, Switzerland, 2020; Volume 158, pp. 131–156.
8. Hwang, G.J.; Sung, H.Y.; Chang, S.C.; Huang, X.C. A fuzzy expert system-based adaptive learning approach to improving students' learning performances by considering affective and cognitive factors. *Comput. Educ.* **2020**, *1*, 100003. [CrossRef]
9. Fodor, K.; Balogh, Z. Sensory Monitoring of Physiological Functions Using IoT Based on a Model in Petri Nets. In *Web Information Systems Engineering—WISE 2021*; Zhang, W.J., Zou, L., Maamar, Z., Chen, L., Eds.; Springer: Cham, Switzerland, 2021; Volume 13081, pp. 435–443.
10. Kurniawan, D.A.; Astalini, A.; Darmaji, D.; Melsayanti, R. Students' attitude towards natural sciences. *Int. J. Eval. Res. Educ.* **2019**, *8*, 455. [CrossRef]
11. Feng, Z.C.; Zhou, Z.J.; Hu, C.H.; Chang, L.L.; Hu, G.Y.; Zhao, F.J. A new belief rule base model with attribute reliability. *IEEE Trans. Fuzzy Syst.* **2019**, *27*, 903–916. [CrossRef]
12. Patlar Akbulut, F.; Perros, H.G.; Shahzad, M. Bimodal affect recognition based on autoregressive hidden Markov models from physiological signals. *Comput. Methods Programs Biomed.* **2020**, *195*, 105571. [CrossRef]
13. Harper, R.; Southern, J. A Bayesian Deep Learning Framework for End-To-End Prediction of Emotion from Heartbeat. *IEEE Trans. Affect. Comput.* **2022**, *13*, 985–991. [CrossRef]
14. Ray, P.; Chakrabarti, A. A Mixed approach of Deep Learning method and Rule-Based method to improve Aspect Level Sentiment Analysis. *Appl. Comput. Inform.* **2022**, *18*, 163–178. [CrossRef]
15. Liu, T.; Gu, X. Opening the "Black Box": Exploring the Interpretability of Artificial Intelligence in Education. *China Educ. Technol.* **2022**, *5*, 82–90.
16. Saganowski, S. Bringing Emotion Recognition Out of the Lab into Real Life: Recent Advances in Sensors and Machine Learning. *Electronics* **2022**, *11*, 496. [CrossRef]
17. Yang, J.B.; Liu, J.; Wang, J.; Sii, H.S.; Wang, H.W. Belief rule-base inference methodology using the evidential reasoning Approach-RIMER. *IEEE Trans. Syst. Man Cybern. Syst.* **2006**, *36*, 266–285. [CrossRef]
18. Zhou, Z.J.; Hu, G.Y.; Hu, C.H.; Wen, C.L.; Chang, L.L. A Survey of Belief Rule-Base Expert System. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *51*, 4944–4958. [CrossRef]
19. Xu, X.J.; Yan, X.P.; Sheng, C.X.; Yuan, C.Q.; Xu, D.L.; Yang, J.B. A Belief Rule-Based Expert System for Fault Diagnosis of Marine Diesel Engines. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *50*, 656–672. [CrossRef]
20. Cao, Y.; Zhou, Z.J.; Hu, C.H.; Tang, S.W.; Wang, J. A new approximate belief rule base expert system for complex system modeling. *Decis. Support Syst.* **2021**, *150*, 113558. [CrossRef]
21. Zhou, Z.J.; Cao, Y.; Hu, G.Y.; Zhang, Y.M.; Tang, S.W.; Chen, Y. New health-state assessment model based on belief rule base with interpretability. *Sci. China Inf. Sci.* **2021**, *64*, 15. [CrossRef]
22. Hossain, M.S.; Rahaman, S.; Mustafa, R.; Andersson, K. A belief rule-based expert system to assess suspicion of acute coronary syndrome (ACS) under uncertainty. *Soft Comput.* **2018**, *22*, 7571–7586. [CrossRef]

23. Hu, G.X.; He, W.; Sun, C.; Zhu, H.L.; Li, K.L.; Jiang, L. Hierarchical belief rule-based model for imbalanced multi-classification. *Expert Syst. Appl.* **2023**, *216*, 119451. [CrossRef]

24. Yang, Y.; Fu, C.; Chen, Y.W.; Xu, D.L.; Yang, S.L. A belief rule based expert system for predicting consumer preference in new product development. *Knowl. Based Syst.* **2016**, *94*, 105–113. [CrossRef]

25. Wang, Y.M.; Yang, L.H.; Fu, Y.G.; Chang, L.L.; Chin, K.S. Dynamic rule adjustment approach for optimizing belief rule-base expert system. *Knowl. Based Syst.* **2016**, *96*, 40–60. [CrossRef]

26. Chang, L.L.; Zhou, Y.; Jiang, J.; Li, M.J.; Zhang, X.H. Structure learning for belief rule base expert system: A comparative study. *Knowl. Based Syst.* **2013**, *39*, 159–172. [CrossRef]

27. Wang, J.; Zhou, Z.J.; Ning, P.Y.; Liu, S.T.; Zhou, X.Y.; Zhao, Y. Inference and analysis of a new evidential reasoning rule-based performance evaluation model. *Eng. Appl. Artif. Intell.* **2023**, *119*, 105789. [CrossRef]

28. Yang, J.B. Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties. *Eur. J. Oper. Res.* **2001**, *131*, 31–61. [CrossRef]

29. Zhou, Z.J.; Hu, G.Y.; Zhang, B.C.; Hu, C.H.; Zhou, Z.G.; Qiao, P.L. A Model for Hidden Behavior Prediction of Complex Systems Based on Belief Rule Base and Power Set. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *48*, 1649–1655. [CrossRef]

30. Goliatt, L.; Yaseen, Z.M. Development of a hybrid computational intelligent model for daily global solar radiation prediction. *Eng. Appl. Artif. Intell.* **2023**, *212*, 118295. [CrossRef]

31. Hansen, N. The CMA evolution strategy: A comparing review. In *Towards a New Evolutionary Computation*; Lozano, J.A., Larrañaga, P., Inza, I., Bengoetxea, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; Volume 192, pp. 75–102.

32. Hu, G.Y. Study on Network Security Situation Awareness Based on Belief Rule Base. Ph.D. Thesis, Harbin University of Science and Technology, Harbin, China, 2016.

33. Martin, M.O.; Mullis, I.V. *TIMSS 2019 Assessment Frameworks*; TIMSS & PIRLS International Study Center, Boston College and International Association for the Evaluation of Educational Achievement (IEA): Amsterdam, The Netherlands, 2017.