

Article

A Classification Model Based on Interval Rule Inference Network with Interpretability

Yunxia Zhang ^{1,2}, Yiming Zhong ^{1,2,*}, Xiaochang Wu ^{3,*} and Jing Bai ^{1,2}¹ School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang 050043, China; zhangyx01@stdu.edu.cn (Y.Z.); baijing66@stdu.edu.cn (J.B.)² Shijiazhuang Key Laboratory of Artificial Intelligence, Shijiazhuang Tiedao University, Shijiazhuang 050043, China³ School of Computer, Guangdong University of Petrochemical Technology, Maoming 525000, China

* Correspondence: 1202210026@student.stdu.edu.cn (Y.Z.); wuxiaochang@gdpuet.edu.cn (X.W.)

Abstract: Interpretability requirements, complex uncertain data processing, and limited training data are characteristics of classification in some real industry applications. The interval belief rule base (IBRB) can deal with various types of uncertainty and provides high interpretability. However, there is a large number of parameters in IBRB, which makes it difficult for experts to accurately set them manually, limiting its application scope. To address this issue, this paper proposes an interval rule inference network (IRIN) with interpretability for classification models to automatically generate IBRB through integrating the ideas of the IBRB and the neural network. Firstly, hybrid data with different types are transformed into an interval belief distribution for automatic generation processing. Secondly, the interval evidence reasoning method is utilized as the inference engine to transfer information ensuring the process's interpretability. Finally, a reasonable IBRB is generated automatically by updating the parameters by employing the learning engine in the neural network. Moreover, the differentiability of the interval evidence reasoning method in the IRIN is proved as a theoretical foundation of the IRIN, and an interpretability analysis of the IRIN's structures is discussed. Experimental results demonstrate that the proposed method possesses high interpretability, enhancing the reliability of classification and maintaining the accuracy. Its application in an actual engineering case illustrates that it is particularly suitable for engineering problems where the explanation of results is a critical requirement.



Received: 5 December 2024

Revised: 3 January 2025

Accepted: 7 January 2025

Published: 10 January 2025

Citation: Zhang, Y.; Zhong, Y.; Wu, X.; Bai, J. A Classification Model Based on Interval Rule Inference Network with Interpretability. *Appl. Sci.* **2025**, *15*, 649. <https://doi.org/10.3390/app15020649>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: interval rule inference network; rule inference network; interval belief rule base; interval uncertainty; interpretability

1. Introduction

Classification problems have long been a key component of machine learning in many fields. In view of the nonlinearity, complexity, and uncertainty of the real world, a variety of classification methods have been proposed that perform well in solving classification problems. However, many classification methods, especially neural networks, achieve superior performance at the expense of low interpretability. In recent years, there has been a surge of interest in interpretable machine learning techniques [1–4].

To address the opacity issue in the decision making of many classification methods, researchers have begun to explore integrating expert knowledge into classifiers [5]. Rule-based classification systems, which integrate expert knowledge and provide high interpretability, have become a prominent artificial intelligence technology [6]. However,

when designing and implementing a rule-based classification system, due to the diversity of input data and the complex uncertainty of the real world, the system will inevitably face challenges such as fuzziness, incompleteness, and uncertainty [7,8].

To tackle these challenges, Yang et al. [9] proposed the belief-rule-based inference methodology using the evidential reasoning approach (RIMER). This method establishes rules using belief distributions to effectively handle various types of uncertain knowledge. Subsequently, Liu et al. [10] further expanded the BRB and proposed the extended belief rule base (EBRB). The EBRB embeds belief degrees in antecedent attributes as well as the results of each rule to handle uncertainty more efficiently. However, neither the EBRB nor the BRB consider interval uncertainty, so it cannot effectively handle interval uncertainty caused by interval values, interval belief degrees, and interval reference values.

To solve this problem, some researchers have conducted various studies in recent years. Aiming at the interval uncertainty caused by interval reference values, AbuDaha et al. [11] introduced the generalized belief rule base, where grade intervals are used to express consequents, effectively solving the interval uncertainty problem caused by the reference value of the antecedent attribute. Aiming at the interval uncertainty caused by interval values and interval belief degree, Wang et al. [12] extended the ER method to the interval ER method, which effectively dealt with uncertainty, including the interval belief degree and interval value. Gao et al. [7] introduced an interval belief rule-based system (IBRBS) that considers interval data, interval belief degrees, and interval reference values, addressing the overall interval uncertainty. Zhu et al. [13] proposed an interval-valued belief rule-based inference method using the evidential reasoning approach (IRIMER). The IRIMER embeds interval-valued belief degrees in both the antecedents and consequents of each rule, allowing it to represent uncertain information or knowledge more flexibly and reasonably than previous belief rule bases. However, although these methods can effectively deal with interval uncertainty, these systems are all set by experts manually. When the data scales or data dimensions increase, there are a large number of parameters in the IBRB system, which inevitably increases the number of parameters. The complexity creates the risk of subjective fallacy. Therefore, determining how to determine these parameters has become a key issue in IBRB research.

To date, research on the automatic generation of IBRB is still insufficient. In other rule-based systems, there are some methods for automatic generation, which provide research ideas for building methods for automatically generating IBRBs. Liu et al. [14] proposed a student performance prediction model based on a belief rule base with automatic construction, where the accuracy of the model is improved by optimizing the parameters using the P-CMA-ES algorithm. Zhang et al. [15] proposed a new parameter optimization approach based on an improved gray wolf optimizer with interpretability reinforcement, which contributes to the interdisciplinary research on BRB systems and metaheuristic algorithms. Fu et al. [16] proposed a new EBRB system based on the K-nearest neighbor graph index. In the optimization model, all samples are utilized for gradient updates in each training iteration to update the parameters. Huang et al. [17] used a gradient descent algorithm to adjust parameters in the BRB and proposed a rule inference network model to automatically generate a BRB.

To automatically generate an IBRB and break through the limitations of its application, this paper proposes an interpretable interval rule inference network by combining interval belief rule bases with neural networks. This approach effectively handles interval uncertainty, enhances system reliability, and leverages the learning mechanism of neural networks to reduce dependency on experts setting parameters, enabling automatic updates of the interval belief rule base.

The main contributions of this paper include three aspects:

- (1) An interpretable rule inference network (IRIN) for the classification model is proposed. This model automatically generates an interval belief rule base, avoiding the problem that experts cannot accurately set IBRB parameters. It can improve the classification performance as well, extending the application areas compared with the RIN.
- (2) The feedforward process of IRIN is the inference process of the IBRB, and an IBRB is obtained to explain the algorithm results. It ensures the interpretability of the model, improving the research in the field of interpretable machine learning, meeting the interpretability, the intervenible ability, and reliability required in some.
- (3) Experiments on classification tasks are carried out on multiple public datasets. The experimental results show that IRIN has high interpretability and performs better than other BRB-based methods. This model is used in an engineering application, providing strong support for the practical application of IRIN.

2. Related Work

A rule-based system is an interpretability artificial intelligence technology. Rules can manipulate and incorporate new and existing knowledge, and adjust the behavior of the system through rules. At present, rule-based systems are mainly divided into three types: belief rule base (BRB) [9], extended belief rule base (EBRB) [10], and interval belief rule base (IBRB) [13].

The BRB, as an interpretable language model, has the ability to process qualitative and quantitative information under uncertainty and has developed rapidly in recent years. Gao et al. [18] proposed a fast and accurate BRB generation and reduction method, which grouped and combined similar rules in BRBs and removed the redundancy-based belief rule to alleviate the combinatorial explosion problem of BRBs. Yang et al. [19] proposed hierarchical BRB using a multi-layer tree structure (MTS-BRB), which effectively overcomes the combinatorial explosion problem of BRBs by introducing a multi-layer tree structure. Jiao et al. [20] proposed to learn a compact and accurate belief rule base with the help of the decision tree building technique. By introducing a decision tree build technology that jointly implements feature selection and model construction, a compact and accurate BRB was obtained.

On the basis of the BRB, Liu et al. [10] further proposed an extended BRB, namely, EBRB. EBRB not only inherits the advantages of the BRB inference model, but also has an effective and concise modeling method, so it has attracted the attention of many researchers. Yang et al. [5] introduced data envelopment analysis to evaluate the efficiency of each rule in EBRB, effectively reducing the number of EBRB rules. Fu et al. [21] proposed a rule activation method based on nearest neighbor propagation, which significantly improved the accuracy of EBRB inference and efficiency. Ma et al. [22] proposed a rule generation and activation method for the extended belief rule system based on an improved decision tree, and constructed an effective classification model. This model effectively solved the problems of low EBRB inference efficiency and inconsistent rule activation.

However, since neither EBRB nor BRB considers interval uncertainty, they are limited in handling this type of uncertainty. To solve this problem, some researchers proposed a new rule-based system called IBRB. IBRB effectively handles various types of interval uncertainty by embedding interval-valued belief degrees in both the antecedents and consequents of each rule. Currently, IBRB is still in the improvement and development stage, and most of the research focuses on how to deal with the interval uncertainty caused by interval data, interval reference values, and interval belief degree [7,11–13].

It can be found from the above-mentioned related works on rule-based systems that the IBRB is a more flexible and rational rule-based system for expressing uncertain information and knowledge. However, it is a challenge for experts to set multiple parameters manually

in IBRB, especially when the number of rules is large. Hence, the present study focuses on using a method that can automatically generate IBRBs to solve the highly complex problem of manual parameter setting by IBRB experts.

3. Preliminaries

3.1. Interval Belief Rule Base

The IBRB was established with interval belief distributions embedded in both the antecedent and consequent terms of each rule, and is capable of capturing interval uncertainty and incompleteness in an integrated way. The IBRB has multiple interval belief rules, and the k th rule is profiled as follows:

$$\begin{aligned} R_k : & \text{IF } (U_1 \text{ is } \{A_{n,1}^k, [\alpha_{n,1}^{k-}, \alpha_{n,1}^{k+}]\} (n = 1, \dots, j_i)) \wedge \dots \\ & \wedge (U_T \text{ is } \{A_{m,T}^k, [\alpha_{m,T}^{k-}, \alpha_{m,T}^{k+}]\} (m = 1, \dots, j_i)), \\ & \text{THEN } (D_1, [\beta_{1,k}^-, \beta_{1,k}^+]), \dots, (D_{j_D}, [\beta_{j_D,k}^-, \beta_{j_D,k}^+]), \\ & \text{with a rule weight } \theta_k (k = 1, \dots, L) \\ & \text{and antecedent attribute weights } \delta_T (T = 1, \dots, T) \end{aligned} \quad (1)$$

where $(A_{j,i}^k, [\alpha_{j,i}^{k-}, \alpha_{j,i}^{k+}]) (i = 1, \dots, T, j = 1, \dots, j_i)$ is the interval-valued belief distribution of the i th antecedent, $[\alpha_{j,i}^{k-}, \alpha_{j,i}^{k+}]$ is the interval-valued belief degree to which the i th attribute U_i is evaluated to be the referential value $A_{j,i}^k$ and $\alpha_{j,i}^k \in [\alpha_{j,i}^{k-}, \alpha_{j,i}^{k+}]$ such that $\sum_{j=1}^{j_i} \alpha_{j,i}^k \leq 1$; $[\beta_{j,k}^-, \beta_{j,k}^+]$ is the interval-valued belief degree to which D_n is believed to be the consequent of the k th rule.

3.2. Rule Inference Method

The inference process of IBRB can be divided into three steps: input transformation, activation weight calculation, and the rule inference method based on the IER algorithm.

Assume that an input for the interval belief rule-based system is $(x_1, x_2, \dots, x_T)$, $x_i (i = 1, \dots, T)$ is the input for the i th antecedent attribute U_i , and T is the total number of the antecedent attributes involved in all the rules of the IBRB.

Before rule inference, the input can be transformed into a belief structure using a utility based information transformation technique as follows:

$$S(x_i) = \{(A_{i,j}, \alpha_{i,j}), j = 1, \dots, j_i\} \quad (2)$$

where $\alpha_{j,i}^k \in [\alpha_{j,i}^{k-}, \alpha_{j,i}^{k+}]$. For input interval samples, there are usually two different cases.

When $x_i \in [x_i^-, x_i^+]$ is between two reference values:

$$\begin{aligned} \alpha_{i,j}^- &= \frac{\alpha_{ij+1} - x_i^+}{\alpha_{ij+1} - \alpha_{ij}}, & \alpha_{i,j}^+ &= \frac{\alpha_{ij+1} - x_i^-}{\alpha_{ij+1} - \alpha_{ij}} \\ \alpha_{i,j+1}^- &= \frac{x_i^- - \alpha_{ij}}{\alpha_{ij+1} - \alpha_{ij}}, & \alpha_{i,j+1}^+ &= \frac{x_i^+ - \alpha_{ij}}{\alpha_{ij+1} - \alpha_{ij}} \end{aligned} \quad (3)$$

When $x_i \in [x_i^-, x_i^+]$ includes one or more reference values:

$$\begin{aligned} \alpha_{i,j-1}^- &= 0, \alpha_{i,j-1}^+ = I_{j-1,j} \times \frac{x_i^- - \alpha_{i,(j-1)}}{\alpha_{i,j} - \alpha_{i,(j-1)}} \\ \alpha_{i,j}^- &= 0, \alpha_{i,j}^+ = I_{j-1,j} + I_{j,j+1} \\ \alpha_{i,j+1}^- &= 0, \alpha_{i,j+1}^+ = I_{j,j+1} + I_{j+1,j+2} \\ \alpha_{i,j+2}^- &= 0, \alpha_{i,j+2}^+ = I_{j+1,j+2} \times \frac{x_i^+ - \alpha_{i,(j+1)}}{\alpha_{i,j+2} - \alpha_{i,(j+1)}} \end{aligned} \quad (4)$$

where if $x_i \in [x_i^-, x_i^+]$ is between $\alpha_{i,j-1}$ and $\alpha_{i,j}$, then $I_{j-1,j} = 1$, otherwise $I_{j-1,j} = 0$.

Once the input data are converted, the similarity between the i th attribute x_i of the input and U_i of the k th rule is measured, which can be calculated by

$$s_i^k = 1 - \sqrt{\frac{1}{j_i} \sum_{j=1}^{j_i} \frac{(\alpha_{i,j}^{k-} - \alpha_{i,j}^{k-})^2 - (\alpha_{i,j}^{k+} - \alpha_{i,j}^{k+})^2}{2}} \quad (5)$$

Next, the activation weight for the k th belief rule can be calculated as follows:

$$w_k = \frac{\theta_k \times \prod_{i=1}^{T_k} (s_i^k)^{\bar{\delta}_i}}{\sum_{j=1}^L (\theta_k \times \prod_{i=1}^{T_k} (s_i^k)^{\bar{\delta}_i})} \quad (6)$$

with

$$\bar{\delta}_i = \frac{\delta_i}{\max_{i=1,\dots,M} \{\delta_i\}} \quad (7)$$

After calculating the activation weights for each rule, the activated rules are aggregated based on the IER approach to generate the combined belief degree. The calculation is as follows:

$$m_{j,k} = m_k(D_j) \in [m_{j,k}^-, m_{j,k}^+] = [w_k \times \beta_{j,k}^-, w_k \times \beta_{j,k}^+] \quad (8)$$

$$\begin{aligned} \tilde{m}_{D,k} &= \tilde{m}_k(D) \in [\tilde{m}_{D,k}^-, \tilde{m}_{D,k}^+] = [w_k \times \beta_{D,k}^-, w_k \times \beta_{D,k}^+] \\ &= [w_k \times (1 - \sum_{j=1}^{J_D} \beta_{j,k}^+), w_k \times (1 - \sum_{j=1}^{J_D} \beta_{j,k}^-)] \end{aligned} \quad (9)$$

$$\bar{m}_{D,k} = \bar{m}_k(D) = 1 - w_k \quad (10)$$

where $m_{j,k}$ represents the credibility that the k th rule is not assigned to any consequent attribute, and $\tilde{m}_{D,k}$ represents the credibility of the missing reference attribute of the k th rule. The total uncertainty credibility is given by $m_{D,k} = \tilde{m}_{D,k} + \bar{m}_{D,k}$. We synthesize the credibility information of all rules and obtain the final belief result of each consequent attribute:

$$m_j = K[\prod_{k=1}^L (m_{j,k} + \tilde{m}_{D,k} + \bar{m}_{D,k}) - \prod_{k=1}^L (\tilde{m}_{D,k} + \bar{m}_{D,k})] \quad (11)$$

$$\tilde{m}_D = K[\prod_{k=1}^L (\tilde{m}_{D,k} + \bar{m}_{D,k}) - \prod_{k=1}^L (\bar{m}_{D,k})] \quad (12)$$

$$\bar{m}_D = K[\prod_{k=1}^L \bar{m}_{D,k}] \quad (13)$$

$$K = [\sum_{j=1}^{J_D} \prod_{k=1}^L (m_{j,k} + \tilde{m}_{D,k} + \bar{m}_{D,k}) - (J_D - 1) \prod_{k=1}^L (\tilde{m}_{D,k} + \bar{m}_{D,k})]^{-1} \quad (14)$$

$$\max / \min y_i = \frac{m_j}{1 - \bar{m}_D} \quad (15)$$

4. Interval Rule Inference Network

In this section, an interval rule inference network model is constructed. This model combines the interval belief rule base and neural network, which can not only handle a variety of uncertain information or knowledge including interval uncertainty, but also automatically generate an IBRB through learning algorithms. To facilitate the automatic generation of an IBRB, the framework of the proposed IRIN is presented in Figure 1.

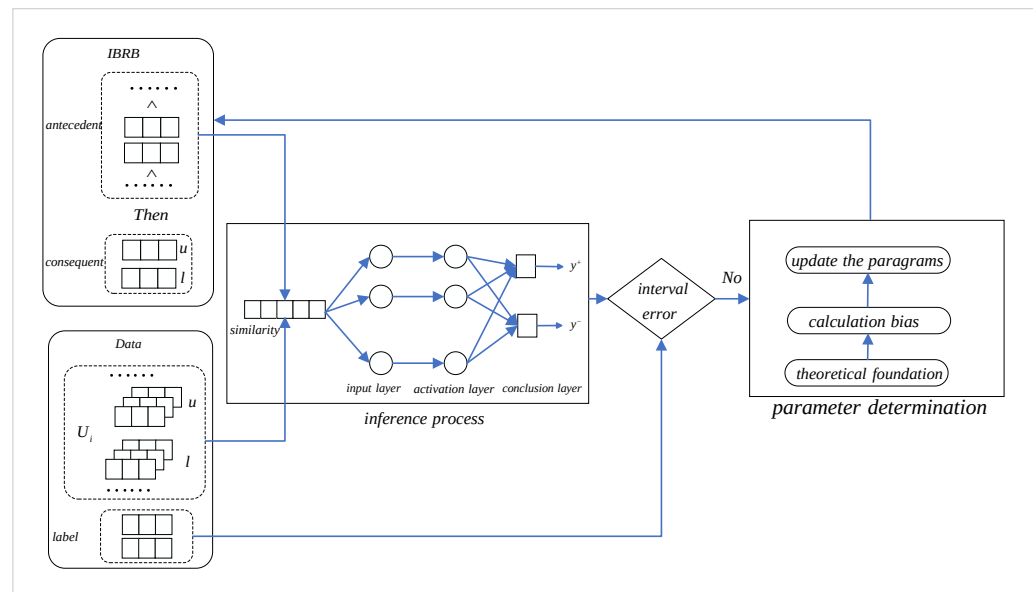


Figure 1. IRIN network framework.

The IRIN is composed of three parts, namely, the IBRB establishment, the inference process, and the parameter determination. The IBRB establishment contains a transformation process where the values of samples of any type are translated into interval belief distributions by statistical methods [23–25], and rule base establishment where an IBRB is constructed associated with the samples. The inference process is divided into three layers, i.e., the input layer, the activation layer, and the conclusion layer.

The similarity measures are input into the IBRB, in which the sample and the rules in IBRB are both represented with interval belief distributions, and the IER-based method is used to reason the information with the interval belief distribution values. The conclusion is obtained containing interval belief distribution information.

In the parameter determination, the IBRB is optimized using gradient descent, where the error of mean square is calculated between the outputs of the inference process and the actual labels, which are all represented as interval belief distribution values. The belief degrees of consequents $\beta_{j,k} (j = 1, \dots, J_D; k = 1, \dots, L)$, $\beta_{j,k} \in [\beta_{j,k}^-, \beta_{j,k}^+]$ should be handled as interval belief distribution, while attribute weights $\delta_i (i = 1, \dots, T)$ and rule weights $\theta_k (1, \dots, L)$ are real values.

4.1. IBRB Establishment

In some real-world classification problems, samples are represented by numerical values or interval values. To establish IBRB, data preparation is necessary, including two steps, interval value conversion and belief distribution conversion. An IBRB is constructed using datasets with interval belief distribution representation obtained by the data preparation.

4.1.1. Data Transformation

In the interval value conversion, numerical values are converted into interval values. This process ensures the equivalence and rationality of the information before and after the data conversion.

Assume the i th sample in the dataset is $x_i = \{x_{1,i}, \dots, x_{t,i}, \dots, x_{T,i}; y_i\}$, in which $x_{t,i}$ is general numerical representation, T denotes the total number of the attribute, y_i represents the category of the i th sample, and $y_i \in \{y_1, \dots, y_{J_D}\}$, where J_D indicates the total number of categories. The method for converting a numerical value into an interval value $x_{t,i} \in [x_{t,i}^-, x_{t,i}^+]$ is as follows:

$$\begin{cases} x_{t,i}^- = x_{t,i} - \text{rand}() \sigma_{t,j} \lambda, \\ x_{t,i}^+ = x_{t,i} + \text{rand}() \sigma_{t,j} \lambda, \end{cases} \quad \text{if } y_i = y_j \quad (16)$$

where λ is the step size, and $\sigma_{t,j}$ is the standard deviation of x_t in the same class y_j .

$$\sigma_{t,j} = \sqrt{\frac{1}{|y_j| - 1} \sum_{x_{t,i} \in y_j} (x_{t,i} - \bar{x}_{t,j})^2} \quad (17)$$

with

$$\bar{x}_{t,j} = \frac{1}{|y_{JD}|} \sum_{x_{t,i} \in y_j} x_{t,i} \quad (18)$$

The category of the sample should be converted to interval values $[y_i^-, y_i^+]$ as follows:

$$\begin{cases} y_i^- = y_i - \frac{1}{2}\gamma \\ y_i^+ = y_i + \frac{1}{2}\gamma \end{cases} \quad (19)$$

where γ is random number and $\gamma \in [0, 1]$.

A given input x_t is transformed into an interval belief distribution of the antecedent attribute U_t as follows:

$$E(x_t) = \{ (A_{tj}, [a_{t,j}^-, a_{t,j}^+]) \mid j = 1, \dots, J \}, t = 1, \dots, T \quad (20)$$

where A_{tj} is the j th reference value of the t th antecedent attribute, and $[a_{t,j}^-, a_{t,j}^+]$ is the interval belief degree to which the input x_t belongs to the referential value A_{tj} . $[a_{t,j}^-, a_{t,j}^+]$ will be obtained in different cases from Equations (3) and (4) according to the nature of the antecedent attribute.

The category of samples can be expressed as the interval belief distribution representation. Assuming that the sample belongs to the i th category, then it can be expressed as

$$\{(D_1, b_1), \dots, (D_{JD}, b_{JD})\} \quad (21)$$

$$\text{where } b_j = [b_j^-, b_j^+], b_j^- = b_j^+ = \begin{cases} 1, j \in i \\ 0, j \notin i \end{cases}$$

4.1.2. IBRB Generation

The sample represented by the above interval belief distribution representation can obtain the IBRB in the form shown in Equation (1); the IBRB is shown in Table 1.

In Table 1, $A_{t,j}^k$ ($k = 1, \dots, L; t = 1, \dots, T; j = 1, \dots, J$) is the j th reference value of the antecedent attribute U_t ($t = 1, \dots, T$) of the k th rule, T is the total number of the antecedent attributes involved in all the rules of the IBRB, J represents the number of reference values, L represents the number of IBRB rules, $\bar{\delta}_t$ represents the attribute weight, $\alpha_{t,j}^k = [\alpha_{t,j}^{k-}, \alpha_{t,j}^{k+}]$ is the interval belief degree belongs to $A_{t,j}^k$. $\beta_{k,JD} = [\beta_{k,JD}^-, \beta_{k,JD}^+]$ the interval-valued belief degree to which D_{JD} is believed to be the consequent of the k th rule, and J_D represents the number of categories.

The antecedent and consequent terms of each IBR in the IBRB are embedded with interval belief distributions, effectively capturing the interval uncertainty associated with

the input samples. As the IBRB generated in Table 1 shows, the k th IBR can be represented as follows:

$$\begin{aligned}
 R_k : & \text{IF } (U_1 \text{ is } \{(A_{1,1}^k, [\alpha_{1,1}^{k-}, \alpha_{1,1}^{k+}]), \dots, (A_{1,J}^k, [\alpha_{1,J}^{k-}, \alpha_{1,J}^{k+}])\}) \wedge \dots \\
 & \wedge (U_T \text{ is } \{(A_{T,1}^k, [\alpha_{T,1}^{k-}, \alpha_{T,1}^{k+}]), \dots, (A_{T,J}^k, [\alpha_{T,J}^{k-}, \alpha_{T,J}^{k+}])\}), \\
 & \text{THEN } (D_1, [\beta_{k,1}^-, \beta_{k,1}^+]), \dots, (D_{J_D}, [\beta_{k,J_D}^-, \beta_{k,J_D}^+]), \\
 & \text{with a rule weight } \theta_k \\
 & \text{and antecedent attribute weights } \delta_T (T = 1, \dots, T).
 \end{aligned} \tag{22}$$

Table 1. Rule generation in IBRB.

Rule	Rule Weight	Antecedent			Consequent $D_1, \dots, D_{J_D}$
		$U_1(\bar{\delta}_1)$	...	$U_T(\bar{\delta}_T)$	
R_1	θ_1	$(A_{1,1}^1, a_{1,1}^1), \dots, (A_{1,J}^1, a_{1,J}^1)$	...	$(A_{T,1}^1, a_{T,1}^1), \dots, (A_{T,J}^1, a_{T,J}^1)$	$\beta_{1,1}, \dots, \beta_{1,J_D}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
R_k	θ_k	$(A_{1,1}^k, a_{1,1}^k), \dots, (A_{1,J}^k, a_{1,J}^k)$	...	$(A_{T,1}^k, a_{T,1}^k), \dots, (A_{T,J}^k, a_{T,J}^k)$	$\beta_{k,1}, \dots, \beta_{k,J_D}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
R_L	θ_L	$(A_{1,1}^L, a_{1,1}^L), \dots, (A_{1,J}^L, a_{1,J}^L)$	...	$(A_{T,1}^L, a_{T,1}^L), \dots, (A_{T,J}^L, a_{T,J}^L)$	$\beta_{L,1}, \dots, \beta_{L,J_D}$

4.2. Inference Process

After the IBRB is generated, an inference process is performed in order to obtain the final conclusion. Different from previous systems based on BRB or EBRB, the input for IBRB is represented by an interval belief distribution representation. In order to effectively capture the interval information of the sample and the rule, the matching degree of the i th sample $x_{t,i}$ and the antecedent attribute U_i^k of the k th rule is assessed by the similarity measure between the two interval belief distributions. The matching degree between the input x_i and the antecedent attribute U_i of the k th rule is calculated by Equation (5).

At the activation layer, the activation weight can describe the activation degree of the IBR, which can be obtained by combining the matching degree obtained above with attribute weight and rule weight. The activation weight is calculated by Equation (6).

As the input for IRIN is interval information, the final conclusion should also contain interval information. At the conclusion layer, the final conclusion is obtained by aggregating the belief degrees of consequents with interval information and the activation weight of each IBR. The IER-based method is used to aggregate the belief degrees of consequents with interval information and the activation weight of each IBR to obtain the final conclusion. The conclusion presented in the form of intervals can make up for the lack of expert knowledge and improve the robustness of the system to a certain extent. The calculation is as follows:

$$y_i = [y_i^-, y_i^+] \tag{23}$$

s.t.

$$\begin{aligned}
 \beta_{j,k}^- \leq \beta_{j,k} \leq \beta_{j,k}^+ \text{ and } \sum_{j=1}^{J_D} \beta_{j,k} &= 1 \\
 \sum_{k=1}^L w_k &= 1
 \end{aligned} \tag{24}$$

where

$$y_i^- = \frac{\mu^- \cdot [\prod_{k=1}^L (w_k \beta_{j,k}^- + 1 - w_k) - \prod_{k=1}^L (1 - w_k)]}{1 - \mu^- \cdot [\prod_{k=1}^L (1 - w_k)]} \quad (25)$$

$$y_i^+ = \frac{\mu^+ \cdot [\prod_{k=1}^L (w_k \beta_{j,k}^+ + 1 - w_k) - \prod_{k=1}^L (1 - w_k)]}{1 - \mu^+ \cdot [\prod_{k=1}^L (1 - w_k)]} \quad (26)$$

$$\mu^- = [\sum_{j=1}^{J_D} \prod_{k=1}^L (w_k \beta_{j,k}^- + 1 - w_k) - (J_D - 1) \prod_{k=1}^L (1 - w_k)]^{-1} \quad (27)$$

$$\mu^+ = [\sum_{j=1}^{J_D} \prod_{k=1}^L (w_k \beta_{j,k}^+ + 1 - w_k) - (J_D - 1) \prod_{k=1}^L (1 - w_k)]^{-1} \quad (28)$$

For classification problems, the final classification result of the sample is the category with the highest average belief degree.

$$c_j = \frac{y_j^- + y_j^+}{2} \quad (29)$$

where $c_j (j = 1, \dots, J_D)$ is the average belief degree.

$$f(x) = D_n, n = \arg \max_{j=1, \dots, J_D} \{c_j\} \quad (30)$$

4.3. Parameter Determination

To automatically generate a reasonable IBRB, parameters should be determined through the learning algorithm of IRIN, including the attribute weights, rule weights, and the belief degrees of consequents. The gradient descent algorithm is used in the parameter learning process. It avoids the problem of the subjectivity and difficulty of parameter setting manually by experts when the number of parameters is too large, and improves the classification effect of IRIN. Thereby, IRIN can be applied more widely in various areas.

The parameter determination flowchart of IBRB is shown in Figure 2. In order to obtain a reasonable IBRB, the error function E_i is used to calculate the difference between the final conclusion obtained by the inference process and the actual label of the input sample. If this difference is small enough, a reasonable IBRB is considered to have been generated. Otherwise, the belief degree of consequents, rule weights, and attribute weights should be updated continuously.

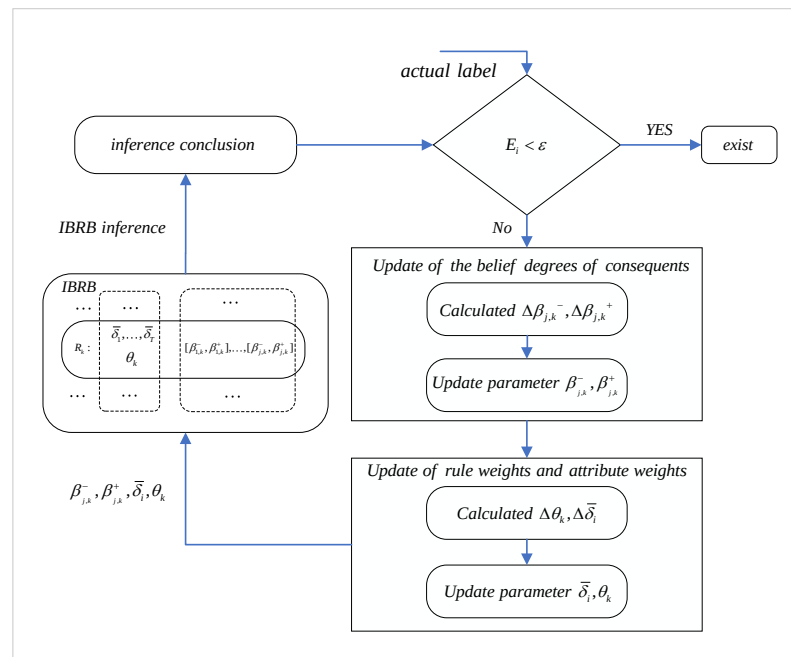


Figure 2. Flowchart of IRIN parameter determination.

4.3.1. The Theoretical Foundations of IRIN

The traditional gradient descent algorithm is adopted to adjust the parameters in IRIN. In order to update the parameters by partial derivation in different directions for different parameters, it is necessary to prove the differentiability of the approximate reasoning algorithm based on IER in IRIN.

Theorem 1. If $\sum_{j=1}^{J_D} \prod_{k=1}^L \beta_{j,k}^+ \neq 0$, $\sum_{j=1}^{J_D} \prod_{k=1}^L \beta_{j,k}^- \neq 0$ and $0 < w_k < 1$, $0 \leq \beta_{j,k}^- \leq \beta_{j,k}^+ \leq 1$, the partial derivatives of the overall combined degree y_i in Equations (25)–(26) to $\beta_{j,k}^-$, $\beta_{j,k}^+$ and w_k exist.

Proof of Theorem 1. Let

$$f_j = [f_j^-, f_j^+] = [\prod_{k=1}^L (w_k \beta_{j,k}^- + 1 - w_k), \prod_{k=1}^L (w_k \beta_{j,k}^+ + 1 - w_k)] \quad (31)$$

$$g = \prod_{k=1}^L (1 - w_k) \quad (32)$$

It can be seen from the differentiability theorem of functions that the partial differentiable of f_j with respect to $\beta_{j,k}^-$, $\beta_{j,k}^+$ and w_k exist, and g is partially differentiable to w_k .

Let

$$\begin{aligned} \mu_0 = [\mu_0^-, \mu_0^+] &= [\sum_{j=1}^{J_D} \prod_{k=1}^L (w_k \beta_{j,k}^- + 1 - w_k) - (J_D - 1) \prod_{k=1}^L (1 - w_k), \\ &\sum_{j=1}^{J_D} \prod_{k=1}^L (w_k \beta_{j,k}^+ + 1 - w_k) - (J_D - 1) \prod_{k=1}^L (1 - w_k)] \end{aligned} \quad (33)$$

since,

$$\begin{aligned}\mu_0^- &= \sum_{j=1}^{J_D} \prod_{k=1}^L (w_k \beta_{j,k}^- + 1 - w_k) - (J_D - 1) \prod_{k=1}^L (1 - w_k) \\ &> \mu_0^- = \sum_{j=1}^{J_D} \prod_{k=1}^L (1 - w_k) - (J_D - 1) \prod_{k=1}^L (1 - w_k) \\ &= J_D \sum_{j=1}^{J_D} (1 - w_k) - (J_D - 1) \prod_{k=1}^L (1 - w_k) \\ &= \prod_{k=1}^L (1 - w_k) > 0\end{aligned}\quad (34)$$

$$\begin{aligned}\mu_0^+ &= \sum_{j=1}^{J_D} \prod_{k=1}^L (w_k \beta_{j,k}^+ + 1 - w_k) - (J_D - 1) \prod_{k=1}^L (1 - w_k) \\ &> \mu_0^+ = \sum_{j=1}^{J_D} \prod_{k=1}^L (1 - w_k) - (J_D - 1) \prod_{k=1}^L (1 - w_k) \\ &= J_D \sum_{j=1}^{J_D} (1 - w_k) - (J_D - 1) \prod_{k=1}^L (1 - w_k) \\ &= \prod_{k=1}^L (1 - w_k) > 0\end{aligned}\quad (35)$$

Hence, μ_0^- and μ_0^+ are partially differentiable with respect to $\beta_{j,k}^-$, $\beta_{j,k}^+$, and w_k .

And since $\mu = \frac{1}{\mu_0}$, μ is partially differentiable with respect to $\beta_{j,k}^-$, $\beta_{j,k}^+$, and w_k .

Hence, $y_j = \frac{\mu \cdot [f_j - g]}{1 - \mu \cdot g}$ is partially differentiable with respect to $\beta_{j,k}^-$, $\beta_{j,k}^+$, and w_k . $\square$

Theorem 2. If $\sum_{k=1}^L (\theta_k \times \prod_{i=1}^{T_k} (s_i^k)^{\bar{\delta}_i}) \neq 0$, the partial derivatives of activation weight w_k in Equation (5) to rule weight θ_k and attribute weight $\bar{\delta}_i$ exist.

Proof of Theorem 2. Let

$$z = \prod_{i=1}^{T_k} (s_i^k)^{\bar{\delta}_i} \quad (36)$$

z is an exponential function about $\bar{\delta}_i$, so z is a partial differentiable to .

$$w_k = \frac{\theta_k \cdot z}{\sum_{j=1}^L (\theta_k \cdot z)} \quad (37)$$

It is obvious that θ_k, z is differentiable and $\sum_{k=1}^L (\theta_k \cdot z) \neq 0$. Hence, w_k is a partially differentiable to $\bar{\delta}_i, z$. From above, w_k is partially differentiable to $\bar{\delta}_i, \theta_k$.

From Theorems 1 and 2, y_i is a partial derivative about $\bar{\delta}_i, \theta_k$. $\square$

4.3.2. Parameter Training

To obtain a reasonable IBRB and build a high-performance classification system, IRIN updates parameters such as $\beta_{j,k}^-$, $\beta_{j,k}^+$, θ_k , $\bar{\delta}_i$, and the gradient descent algorithm is adopted.

1. Calculate the discrepancy:

A given input i th sample is denoted as $x_i = \{x_{1,i}, \dots, x_{t,i}, \dots, x_{T,i}; y_i\}$, and the conclusion obtain by inference from the IBRB is denoted as $\hat{y}_i = [\hat{y}_i^-, \hat{y}_i^+]$, where $\hat{y}_i^- = \{\hat{y}_1^{(i)-}, \dots, \hat{y}_{J_D}^{(i)-}\}$, $\hat{y}_i^+ = \{\hat{y}_1^{(i)+}, \dots, \hat{y}_{J_D}^{(i)+}\}$, x_i is the i th sample and $y_i =$

$[\{y_1^{(i)-}, \dots, y_{J_D}^{(i)-}\}, \{y_1^{(i)+}, \dots, y_{J_D}^{(i)+}\}]$. The discrepancy between the $\hat{y}_i$ and y_i is calculated according to the following loss function.

$$E_i = \frac{1}{J_D} \sum_{j=1}^{J_D} \frac{(\hat{y}_j^- - y_j^-)^2 + (\hat{y}_j^+ - y_j^+)^2}{2} \quad (38)$$

2. Update of the belief degrees of consequents:

The belief degrees of consequents can be updated as

$$\beta_{j,k}^- = \beta_{j,k}^- + \Delta\beta_{j,k}^- (j = 1, \dots, J_D, k = 1, \dots, L) \quad (39)$$

$$\beta_{j,k}^+ = \beta_{j,k}^+ + \Delta\beta_{j,k}^+ (j = 1, \dots, J_D, k = 1, \dots, L) \quad (40)$$

The belief degrees of consequents biases can be updated as

$$\Delta\beta_{j,k}^- = -\eta \frac{\partial E_i}{\partial y_i^-} \quad (41)$$

$$\Delta\beta_{j,k}^+ = -\eta \frac{\partial E_i}{\partial y_i^+} \quad (42)$$

where η is the learning rate, and $\frac{\partial E_i^-}{\partial y_i^-}, \frac{\partial E_i^+}{\partial y_i^+}$ are the gradients.

The partial derivative $\frac{\partial E_i^-}{\partial y_i^-}, \frac{\partial E_i^+}{\partial y_i^+}$ is calculated as follows:

$$\begin{aligned} \frac{\partial E_i}{\partial \beta_{j,k}^-} &= \sum_{j=1}^{J_D} \frac{\partial E_i}{\partial y_j^-} \cdot \frac{\partial y_j^-}{\partial \beta_{j,k}^-} \\ \frac{\partial E_i}{\partial \beta_{j,k}^+} &= \sum_{j=1}^{J_D} \frac{\partial E_i}{\partial y_j^+} \cdot \frac{\partial y_j^+}{\partial \beta_{j,k}^+} \end{aligned} \quad (43)$$

$$\begin{aligned} \frac{\partial E_i}{\partial y_j^-} &= \frac{1}{J_D} (\hat{y}_j^- - y_j^-) \\ \frac{\partial E_i}{\partial y_j^+} &= \frac{1}{J_D} (\hat{y}_j^+ - y_j^+) \end{aligned} \quad (44)$$

$$\begin{aligned} \frac{\partial y_j^-}{\partial \beta_{j,k}^-} &= \frac{(\mu_{\beta}^{-'} f_j^- + \mu^- f_{j\beta}^{-'}) \cdot g^- + (1 - g^-) \mu_{\beta}^{-'} f_j^-}{g^{-2}} \\ \frac{\partial y_j^+}{\partial \beta_{j,k}^+} &= \frac{(\mu_{\beta}^{+'} f_j^+ + \mu^+ f_{j\beta}^{+'}) \cdot g^+ + (1 - g^+) \mu_{\beta}^{+'} f_j^+}{g^{+2}} \end{aligned} \quad (45)$$

where

$$f_j = \left[\prod_{k=1}^L (w_k \times \beta_{j,k}^- + 1 - w_k) - \prod_{k=1}^L 1 - w_k, \prod_{k=1}^L (w_k \times \beta_{j,k}^+ + 1 - w_k) - \prod_{k=1}^L 1 - w_k \right] \quad (46)$$

$$g = \left[1 - \mu^- \prod_{k=1}^L (1 - w_k), 1 - \mu^+ \prod_{k=1}^L (1 - w_k) \right] \quad (47)$$

$$s_k = \left[\prod_{l=1, l \neq k}^L w_l \times \beta_{j,l}^- + 1 - w_l, \prod_{l=1, l \neq k}^L w_l \times \beta_{j,l}^+ + 1 - w_l \right] \quad (48)$$

$$h_k = \prod_{l=1, l \neq k}^L (1 - w_l) \quad (49)$$

$$\text{Let } \mu_0^- = \frac{1}{\mu^-}, \mu_0^+ = \frac{1}{\mu^+}$$

$$\mu_{\beta}^{-'} = \frac{\partial \mu_{\beta}^{-'}}{\partial \beta_{j,k}^-} = -\frac{w_k \cdot s_k^-}{\mu_0^{-2}}, \mu_{\beta}^{+'} = \frac{\partial \mu_{\beta}^{+'}}{\partial \beta_{j,k}^+} = -\frac{w_k \cdot s_k^+}{\mu_0^{+2}} \quad (50)$$

$$\begin{aligned} f_{j\beta}^{-'} &= \frac{\partial f_j}{\partial \beta_{j,k}^-} = w_k \cdot s_k^- \\ f_{j\beta}^{+'} &= \frac{\partial f_j}{\partial \beta_{j,k}^+} = w_k \cdot s_k^+ \end{aligned} \quad (51)$$

3. Update the rule weights and attribute weights:

The rule weights and attribute weights can be updated as

$$\begin{aligned} \theta_k &= \theta_k + \Delta \theta_k (k = 1, \dots, L) \\ \bar{\delta}_i &= \bar{\delta}_i + \Delta \bar{\delta}_i (i = 1, \dots, T) \end{aligned} \quad (52)$$

The rule weights and attribute weights biases can be updated as

$$\begin{aligned} \Delta \theta_k &= -\eta \frac{\partial E_i}{\partial \theta_k} \\ \Delta \bar{\delta}_i &= -\eta \frac{\partial E_i}{\partial \bar{\delta}_i} \end{aligned} \quad (53)$$

From Theorem 1 and Theorem 2, $\frac{\partial E_i^-}{\partial w_k}, \frac{\partial E_i^+}{\partial w_k}, \frac{\partial w_k}{\partial \theta_k}, \frac{\partial w_k}{\partial \bar{\delta}_i}$ exists. According to the inference process, the activation weight of the rule is a function with respect to all the attribute weights, and the output function is a complex function with respect to all the activation weights, bringing high complexity of parameter learning. In order to reduce the complexity, the “pseudo-gradient” will be used instead of the complex “true gradient”, that is, the partial derivative $\frac{\partial \Delta w_k}{\partial \bar{\delta}_i}, \frac{\partial \Delta w_k}{\partial \theta_k}$ will be used instead of $\frac{\partial E_i}{\partial \bar{\delta}_i}, \frac{\partial E_i}{\partial \theta_k}$ as the gradient. Therefore,

$$\begin{aligned} \Delta \theta_k &= -\eta_2 \frac{\partial \Delta w_k}{\partial \theta_k} = -\eta_1 \eta_2 \frac{\partial E_i}{\partial w_k} \cdot \frac{\partial w_k}{\partial \theta_k} \\ \Delta \bar{\delta}_i &= -\eta \frac{\partial \Delta w_k}{\partial \bar{\delta}_i} = -\eta_1 \eta_2 \frac{\partial E_i}{\partial w_k} \cdot \frac{\partial w_k}{\partial \bar{\delta}_i} \end{aligned} \quad (54)$$

The partial derivative $\frac{\partial E_i}{\partial \bar{\delta}_i}, \frac{\partial E_i}{\partial \theta_k}$ is calculated as follows:

$$\frac{\partial E_i}{\partial \theta_k} = \frac{1}{2} \left(\frac{\partial E_i}{\partial y_i^+} \cdot \frac{\partial y_i^+}{\partial w_k} + \frac{\partial E_i}{\partial y_i^-} \cdot \frac{\partial y_i^-}{\partial w_k} \right) \cdot \frac{\partial w_k}{\partial \theta_k} \quad (55)$$

$$\frac{\partial E_i}{\partial \bar{\delta}_i} = \frac{1}{2} \left(\frac{\partial E_i}{\partial y_i^+} \cdot \frac{\partial y_i^+}{\partial w_k} + \frac{\partial E_i}{\partial y_i^-} \cdot \frac{\partial y_i^-}{\partial w_k} \right) \cdot \frac{\partial w_k}{\partial \bar{\delta}_i} \quad (56)$$

where

$$\frac{\partial y}{\partial w_k} = \frac{(\mu_w' f_n + \mu f_{nw}') \cdot g - g' \mu f_n}{g^2} \quad (57)$$

$$\begin{aligned} \mu_w' &= \left[\frac{\partial \mu^{-'}}{\partial w_k}, \frac{\partial \mu^{+'}}{\partial w_k} \right] \\ &= \left[\frac{\sum_j^{J_D} (\beta_{j,k}^- - 1) \cdot s_k^- + (J_D - 1) \cdot h_k}{\mu_0^{-2}}, \right. \\ &\quad \left. \frac{\sum_j^{J_D} (\beta_{j,k}^+ - 1) \cdot s_k^+ + (J_D - 1) \cdot h_k}{\mu_0^{+2}} \right] \end{aligned} \quad (58)$$

$$f_{jw} = [f_{jw}^-, f_{jw}^+] = [(\beta_{j,k}^- - 1) \cdot s_k^- + h_k, (\beta_{j,k}^+ - 1) \cdot s_k^+ + h_k] \quad (59)$$

$$g' = \left[\frac{-\mu_w^-(1-g) + \mu^{2'} \cdot h_k}{\mu^-}, \frac{-\mu_w^+(1-g) + \mu^{2'} \cdot h_k}{\mu^+} \right] \quad (60)$$

Let

$$z_k = \prod_{i=1}^{T_k} (s_i^k)^{\delta_i} \quad (61)$$

If $l = k$

$$\frac{\partial w_k}{\partial \theta_i} = \frac{\theta_k \cdot \sum_{i=1}^{T_k} \theta_i z_i - \theta_k \cdot (\theta_i z_k)}{\left(\sum_{i=1}^{T_k} \theta_i z_i \right)^2} \quad (62)$$

$$\frac{\partial w_k}{\partial \theta_i} = \frac{z_k \cdot \sum_{i=1, i \neq k}^{T_k} \theta_i z_i}{\left(\sum_{i=1}^{T_k} \theta_i z_i \right)^2} \quad (63)$$

$$\begin{aligned} \frac{\partial w_k}{\partial \delta_i} &= \frac{\partial w_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial \delta_i} = \frac{z_k \sum_{i=1}^{T_k} \theta_i z_i - \left(\sum_{i=1}^{T_k} \theta_i z_i \right)' \cdot (\theta_k z_k)}{\left(\sum_{i=1}^{T_k} \theta_i z_i \right)^2} \cdot \ln s_i^k \cdot \prod_{i=1}^{T_k} (s_i^k)^2 \\ &= \frac{z_k \sum_{i=1, i \neq k}^{T_k} \theta_i z_i}{\left(\sum_{i=1}^{T_k} \theta_i z_i \right)^2} \cdot \ln s_i^k \cdot \prod_{i=1}^{T_k} (s_i^k)^2 \end{aligned} \quad (64)$$

If $l \neq k$

$$\frac{\partial w_k}{\partial \theta_i} = -\theta_k z_k \frac{\left(\sum_{j=1, j \neq k}^{T_k} \theta_j z_j \right)'}{\left(\sum_{j=1}^{T_k} \theta_j z_j \right)^2} = -\frac{\theta_k z_k z_i}{\left(\sum_{i=1}^{T_k} \theta_i z_i \right)^2} \quad (65)$$

$$\begin{aligned} \frac{\partial w_k}{\partial \delta_i} &= \frac{\partial w_k}{\partial z_k} \cdot \frac{\partial z_k}{\partial \delta_i} = -\theta_k z_k \frac{\left(\sum_{i=1}^{T_k} \theta_i z_i \right)'}{\left(\sum_{i=1}^{T_k} \theta_i z_i \right)^2} \cdot \ln s_i^k \cdot \prod_{i=1}^{T_k} (s_i^k)^2 \\ &= -\frac{\theta_i z_k z_i}{\left(\sum_{i=1}^{T_k} \theta_i z_i \right)^2} \cdot \ln s_i^k \cdot \prod_{i=1}^{T_k} (s_i^k)^2 \end{aligned} \quad (66)$$

4.4. Algorithm Description

An IRIN that can automatically generate IBR and effectively handle interval uncertainty was constructed to avoid the limitations of parameters given by experts in the above. The algorithm describes the construction process of IRIN. The construction steps of IRIN are as follows:

- Step 1: Normalize the dataset and convert numeric data to interval data.
- Step 2: Use Equations (3) and (4) to transform the interval data into an interval belief distribution according the antecedent attribute.
- Step 3: Refer to Section 4.2 to obtain IBRB inference conclusion $\hat{y}$.
- Step 4: Compute the difference between the IBRB inference conclusion $\hat{y}$ and the actual conclusion y using Equation (38). If the difference is greater than the expected error $\varepsilon_{\min}$, update the parameters by referring to Section 4.3 and return to Step 3.

Step 5: Repeat Step 5 until the difference is less than the expected error $\varepsilon_{\min}$ or the training data are greater than the preset maximum training time $\lambda_{\max}$.

The IRIN algorithm is shown in Algorithm 1.

Algorithm 1: IRIN algorithm.

Input: A training data set $\{(A_i, y_i)\}$, where A_i is the attribute and y_i is the category; Expected error $\varepsilon_{\min}$; Maximum training iterations $\lambda_{\max}$.

Output: IBRB.

```

1 Normalize the dataset;
2 while  $x_i \in \{(A_i, y_i)\}$  do
3   for  $x_{t,i}$  in  $x_i$  do
4     if  $x_{t,i}$  is general numerical representation then
5       Convert  $x_{t,i}$  to  $[x_{t,i}^-, x_{t,i}^+]$  using Equation (16);
6 Transform samples into interval belief distribution representation using Equations (3) and (4);
7 Generate the initial IBRB;
8 Initialize attribute weights  $\bar{\delta}_i (i = 1, \dots, T)$ , rule weights  $\bar{\delta}_i (i = 1, \dots, T)$ , and belief degrees of
   consequents  $\beta_{j,k} \in [\beta_{j,k}^-, \beta_{j,k}^+] (j = 1, \dots, J_D, k = 1, \dots, L)$ ;
9 while  $\varepsilon > \varepsilon_{\min} \wedge \lambda < \lambda_{\max}$  do
10  for  $x_i$  do
11    Compute the individual matching degree using Equation (5);
12    Compute the activation weight using Equation (6);
13    Compute  $\hat{y}_i^- = \{\hat{y}_1^{(i)-}, \dots, \hat{y}_{J_D}^{(i)-}\}$  and  $\hat{y}_i^+ = \{\hat{y}_1^{(i)+}, \dots, \hat{y}_{J_D}^{(i)+}\}$  using Equations (23)–(28);
14    Compute  $c_i = \{\hat{y}_1^{(i)}, \dots, \hat{y}_j^{(i)}\}$  using Equation (29);
15    Compute  $E_i$  using Equation (38);
16    Compute bias by Equations (41), (42) and (53);
17    Update parameters by Equations (39), (40) and (52);
18  Update  $\varepsilon = \text{sum}(E_i) / M$ ;
19  Increment  $\lambda = \lambda + 1$ ;
20 return IBRB;
```

5. Interpretability Analysis of IRIN Structures

In this section, the interpretability of the IRIN classification model constructed in Section 4 will be introduced. To alleviate the problem of the “black box” mechanism, IRIN’s classification model should be interpretable. The authors of [26,27] proposed a general interpretability criterion. Through the analysis of the general interpretability criterion, the constructed IRIN interpretability criterion is as follows.

- Criterion 1. The reference value of the attribute can be effectively distinguished.
When constructing an IRIN, each attribute is assigned at least one reference value with a different meaning. In the present work, the set of referential value of each attribute is $A = (L, 0.0), (M, 0.5), (H, 1.0)$, where H represents “high”, M represents “medium”, and L represents “low”.
- Criterion 2. A complete membership function.
IRIN can convert data with any type of representation into the corresponding interval data, where each datum can match at least one reference value, and at least one rule can be activated in rule inference.
- Criterion 3. The structures and parameters have actual meaning.
As shown in Table 1, all the parameters of the IBRB automatically generated by IRIN have actual meaning, so that is able to deduce a reasonable causal relationship.
- Criterion 4. Standardization of the matching degree.
The standardization of the matching degree helps the understandability of language terms and provides guidance for the classification of IRIN.

- Criterion 5. Reasonable information transformation.
Data in any type or hybrid type need to be converted into interval belief distribution at the data transformation stage. Note that the boundaries of the interval belief distribution obtained by Equations 3 and 4 are not independent interval belief structures and meet the requirement of normalization. Therefore, the data are equivalently expressed in the interval belief structures.
- Criterion 6. A transparent inference engine.
The feedforward process of IRIN is the inference process of IBRB, which uses IER as the inference engine to ensure the interpretability of the rule base in the whole inference process.
- Criterion 7. The simplicity of the rule base.
The numbers of rules, parameters, and reference values for IRIN are moderate to improve the readability of IBRB.

These criteria guarantee the interpretability of IRIN classification and provide a reference for establishing an interpretable classification model.

6. Experiment and Analysis

The datasets used in this paper are described before the experiment. The experiments are divided into three parts. The first part analyzes the interpretability and intervention ability of IRIN, the second part validates the classification performance by comparing the proposed method with other BRB systems, and the third part shows the actual value through applying IRIN to actual engineering.

The experimental environment of this paper is Intel(R) Core(TM) i5-8300H CPU @2.3 GHz (4 CPUs) 24 GB memory, with the Windows 11 operating system. The algorithm in this paper is implemented using the Python 3.7 programming language.

6.1. Experimental Datasets

To verify the effectiveness of the proposed method in this paper, the experiments are conducted on the general datasets from the UC Irvine Machine Learning Repository, with detailed information available at <https://archive.ics.uci.edu/>; accessed on 2 January 2025. The basic information of the datasets used in this paper is presented in Table 2.

In certain engineering domains, historical data for training is not readily available, while the samples for classification prediction are various, such as the classification of mechanical properties of coarse aggregate in the following. Therefore, this paper employs the inverse five-fold cross-validation for experimentation, ensuring the accuracy and reliability of the model in practical applications.

Table 2. Datasets information.

Dataset	Samples	Number of Attributes	Number of Classes
Iris	150	4	3
Seeds	209	7	3
Ecoli	336	7	8
Glass	214	9	7
Haberman	305	3	2
Bupa	344	6	2
Diabetes	768	8	2

6.2. Interpretability Analysis of IRIN

Experts or decision-makers can intervene in IBRB automatically generated by IRIN based on expert knowledge or experience to modify the results to meet different task needs. In order to illustrate the interpretability and the intervention ability, the IBRB with respect to the Iris dataset automatically generated by IRIN is taken as an example, shown in Table 3.

The Iris dataset contains four attributes, namely, sepal length, sepal width, petal length, and petal width, and three classes, where each class refers to a type of iris plant as “Setosa”, “Versicolour”, or “Virginica”. In Table 3, A_1 , A_2 , A_3 , and A_4 represent the antecedent attribute weights, 0.2496, 0.6754, 0.1146, and 0.8475, respectively; D_1 , D_2 , and D_3 represent consequents, which are the class labels; each row is a rule, such as the R_{15} expressed as follows:

$$\begin{aligned}
 R_{15} : & \text{ if the sepal.length is } \{ (L, [0.0002, 0.0080]), (M, [0.9920, 0.9998]), (H, 0.0) \}, \\
 & \text{ AND the sepal.width is } \{ (L, [0.1667, 0.1766]), (M, [0.8234, 0.8333]), (H, 0.0) \}, \\
 & \text{ AND the petal.length is } \{ (L, 0.0), (M, [0.8298, 0.9136]), (H, [0.1667, 0.1701]) \}, \\
 & \text{ AND the petal.width is } \{ (L, 0.0), (M, [0.8333, 0.8412]), (H, [0.1588, 0.1667]) \}, \\
 & \text{ Then the iris category is } \{ (Setosa, [0.2181, 0.3661]), (Versicolour, [0.0595, 0.0935]), \\
 & (Virginica, [0.5404, 0.7224]) \}. \\
 & \text{ with rule weight is } 0.8998 \\
 & \text{ and attribute weight is } \{ (0.2496), (0.6754), (0.1146), (0.8475) \}
 \end{aligned} \tag{67}$$

It represents that if the membership degree of the sepal length at short, average, and long is $[0.0002, 0.008]$, $[0.992, 0.9998]$, 0.0 , and the membership degree of the sepal width at narrow, average, and wide is $[0.1667, 0.1766]$, $[0.8234, 0.8333]$, 0.0 , and the membership degree of the petal length at short, average, and long is 0 , $[0.8298, 0.9136]$, $[0.1667, 0.1701]$, the membership degree of the petal width at narrow, average, and wide is 0.0 , $[0.8333, 0.8412]$, $[0.1588, 0.1667]$, then the membership degrees that this iris category is setosa, versicolour or virginica are $[0.2181, 0.3661]$, $[0.0595, 0.0935]$, and $[0.5404, 0.7224]$, respectively. The rule weight for this rule is 0.8998 , and the antecedent attribute weights of the sepal length, the sepal width, the petal length, and the petal width are 0.2496 , 0.6754 , 0.1146 , and 0.8475 , respectively.

When one or more IBRs in IBRB are inconsistent with expert knowledge information, experts or decision-makers can intervene to modify these IBRs according to expert knowledge or experience, thus correcting the results.

The interpretability of IRIN is illustrated by comparing the consistency of the belief distribution between IRIN and the expert knowledge. More details about the belief distribution can be found in [28]. In Figure 3, the belief distribution of the last 16 rules for IRIN and expert knowledge are depicted. It can be observed that the belief distributions of IBRB are basically consistent with these of expert knowledge, especially rules 28, 27, 20, 19, and 18, which are drawn with pink ellipse. For rules 16, 17, 21, and 22, the differences in their belief distributions compared to expert knowledge may be due to noise interference during the automatic generation of the IBRB by IRIN. However, experts or decision-makers can use their expert knowledge or experience to intervene and modify these rules. Therefore, the belief distribution of the IBRB automatically generated by IRIN can better retain the characteristics of expert knowledge information and effectively handle interval uncertainty, while also improving IRIN’s interpretability and ensuring classification accuracy.

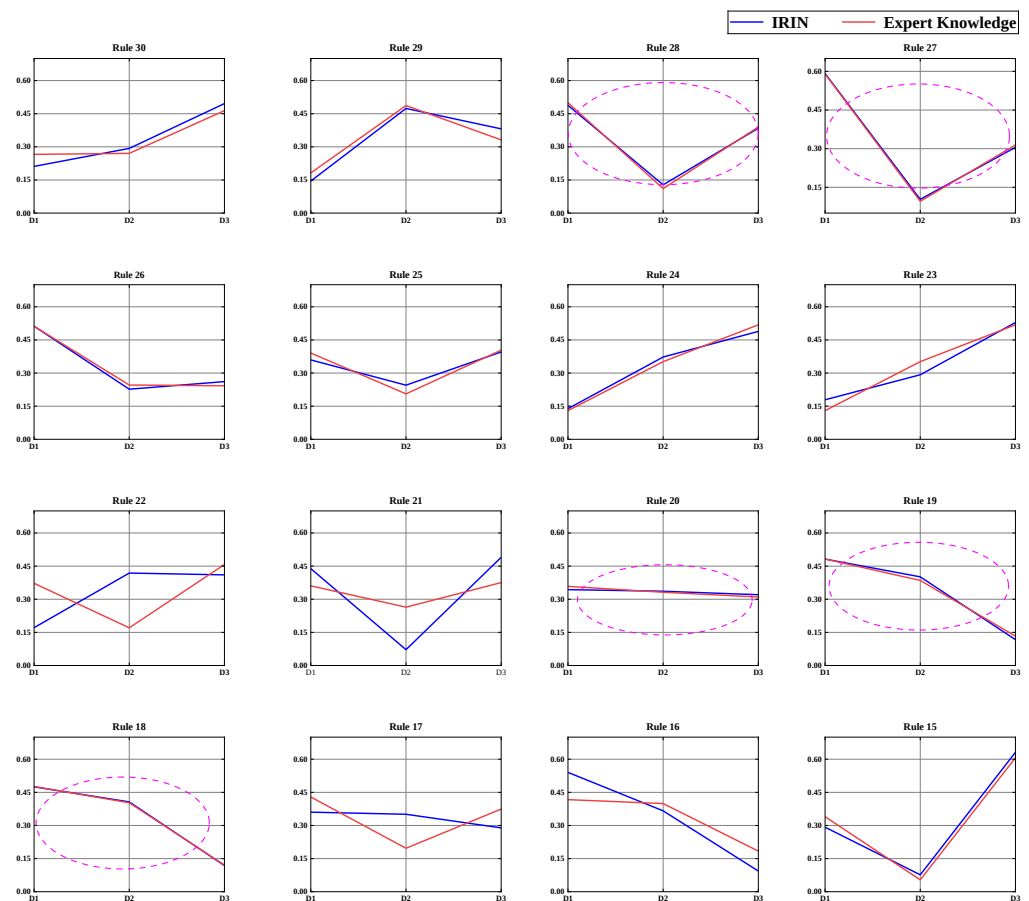


Figure 3. Belief distribution comparison.

Table 3. The IBRB of Iris.

Rule	Rule Weight	Antecedent				Consequent		
		$A_1(0.2496)$	$A_2(0.6754)$	$A_3(0.1146)$	$A_4(0.8475)$	D_1	D_2	D_3
R_1	0.4952	$L([0.3333, 0.3544])$ $M([0.6456, 0.6667])$ $H([0.0, 0.0])$	$L([0.6667, 0.7254])$ $M([0.2747, 0.3333])$ $H([0.0, 0.0])$	$L([0.0899, 0.1509])$ $M([0.8492, 0.9100])$ $H([0.0, 0.0])$	$L([0.1667, 0.2627])$ $M([0.7375, 0.8334])$ $H([0.0, 0.0])$	$[0.2907, 0.3005]$	$[0.3076, 0.3769]$	$[0.3322, 0.3917]$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
R_{15}	0.8998	$L([0.0002, 0.008])$ $M([0.992, 0.9998])$ $H([0.0, 0.0])$	$L([0.1667, 0.1766])$ $M([0.8234, 0.8333])$ $H([0.0, 0.0])$	$L([0.0, 0.0])$ $M([0.6992, 0.7779])$ $H([0.2221, 0.3008])$	$L([0.0, 0.0])$ $M([0.5833, 0.6779])$ $H([0.3221, 0.4167])$	$[0.2181, 0.3661]$	$[0.0595, 0.0935]$	$[0.5404, 0.7224]$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
R_{30}	0.6762	$L([0.0, 0.0])$ $M([0.8333, 0.8446])$ $H([0.1554, 0.1667])$	$L([0.0, 0.1861])$ $M([0.8139, 1.0])$ $H([0.0, 0.0])$	$L([0.0, 0.0])$ $M([0.8298, 0.9136])$ $H([0.1667, 0.1701])$	$L([0.0, 0.0])$ $M([0.8333, 0.8412])$ $H([0.1588, 0.1667])$	$[0.1449, 0.2777]$	$[0.1119, 0.4737]$	$[0.3812, 0.6103]$

6.3. Performance Analysis of IRIN

Parameters in IBRB such as attribute weight, rule weight, and belief degrees of consequents were determined by experts, where the belief degrees of consequents are interval values, complicating the problem of manual parameter setting. Therefore, IRIN effectively avoids this problem by automatically generated an IBRB in which both the antecedent and the consequent term of each rule are embedded in the interval belief distribution.

In order to verify the classification performance of IRIN, the classification accuracy of IRIN is compared with some improved belief rule-based systems.

Table 4 shows the comparison results of classification accuracy with other BRB systems. As can be seen from Table 4, the classification accuracy of IRIN on the Diabetes, Glass, and Ecoli datasets are 75.3%, 76.7%, and 88.1%, respectively, which is 0.9%, 3.8%, and 1.0% higher compared to the second-ranked BRB system. The classification accuracy of the Seeds and Iris datasets is slightly lower than other BRB systems, ranking second and third, respectively. This may be due to the simplicity of the Seeds and Iris datasets, where the class distribution is relatively uniform and the feature dimensions are low. In such a case, the complexity and advantages of the IRIN method are not fully utilized. Nevertheless, the average ranking of IRIN is still better than that of other BRB systems. This fully shows that the IBRB automatically generated using the method proposed can effectively handle interval uncertainty, and the uncertainty can be expressed more carefully by using interval boundaries, so that the system can better distinguish boundary cases and reduce classification errors.

6.4. Application in Mechanical Properties of Coarse Aggregate of Reinforced Recycled Clay Brick

The utilization of recycled brick aggregate as a coarse aggregate in concrete enhances the recycling of waste bricks and reduces resource depletion mixtures. It is meaningful in environmental protection and economy, predicting the classification of the mechanical properties of coarse aggregate of reinforced recycled clay brick in concrete. However, the process requires interpretability and the results should be intervenable, while the history data for training are small. In this study, the proposed IRIN for classification is applied.

Table 4. Comparison of classification results with other BRB systems.

	Iris	Seeds	Ecoli	Glass	Diabetes	Average Rank
EBRB [10]	95.2%(9)	87.1%(9)	81.2%(8)	51.4%(10)	-	36(11)
Yang-EBRB [29]	95.8%(6)	91.2%(7)	59.3%(13)	58.8%(9)	-	35(9)
GSR-BRB [30]	98.6%(1)	94.3%(3)	78.3%(9)	69.1%(7)	72.9%(3)	23(5)
FG-BRB [18]	95.3%(8)	91.2%(7)	75.3%(11)	68.2%(8)	72.7%(4)	38(12)
RD-BRB [18]	98.3%(2)	94.2%(4)	75.8%(10)	68.2%(8)	72.7%(4)	28(8)
DT-BRB [31]	96.0%(5)	94.3%(3)	82.5%(6)	72.9%(4)	-	18(3)
AP-EBRB [21]	95.8%(6)	97.6%(1)	87.1%(2)	50.0%(11)	-	20(4)
DEA-EBRB [32]	95.4%(7)	91.7%(6)	83.3%(5)	69.5%(6)	74.2%(2)	26(7)
IDT-EBRB [22]	97.3%(3)	93.3%(5)	84.6%(4)	70.4%(5)	-	17(2)
MTS-BRB [19]	96.0(5)	87.1%(9)	81.5%(7)	73.4%(2)	68.0%(5)	28(8)
FBDT [20]	96.0%(5)	85.2%(10)	84.6%(4)	65.4%(12)	75.3%(1)	31(9)
RIN [17]	96.6(4)	90.5%(8)	84.8%(3)	60.5%(7)	-	22(5)
IRIN	96.6%(4)	95.2%(2)	88.1%(1)	76.7%(1)	75.3%(1)	9(1)

The data containing 72 samples with five attributions—apparent density ρ , water absorption rate (ω), crushing value (σ), substitution rate (λ_g), and glue–water ratio were collected from Zhongtu International Architectural Design Co. Ltd. (Shijiazhuang, China) and Hebei University of Science and Technology (Shijiazhuang, China) [33]. They are divided into two categories, good mechanical property and bad mechanical property, depending on the compressive strength.

In this experiment, a five-fold cross-validation method was employed to evaluate the performance of the proposed method in identifying the mechanical properties. The evaluation metrics for classifying performance include the following:

- (1) Precision: The proportion of true positive predictions out of all positive predictions.
- (2) Recall: The proportion of true positive predictions out of all actual positives.

- (3) F1 score: The harmonic mean of precision and recall, considering both accuracy and recall metrics. The calculation formula is as follows:

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (68)$$

The comparison of the performance metrics for the seven classification methods is shown in Figure 4. As can be seen from Figure 4, the multilayer perceptron classifier (MLP) and support vector classifier (SVC) have poor performance, while the remaining six methods' classification performance from best to worst are the proposed IRIN, random forest classifier (RFC), decision tree classifier (DTC), RIN, XGBoost (XGB), and K-nearest neighbors classifier (KNC). The proposed IRIN classification method ranks in the top two in all three performance indicators. This shows that IRIN automatically generates reasonable IBRB, which effectively improves the inference effect of IBRB, and then improves the performance of IRIN.

The precision, recall, and F1 score values for the seven classification methods are presented in Table 5. As can be seen from Table 5, the precision of the IRIN model is slightly lower than that of the DTC classifier and RFC classifier by 1.2 percentage points, but IRIN is higher than other classifiers in terms of recall rate and F1 score. Despite the slightly lower precision compared to DTC and RFC classifiers, IRIN's higher recall and F1 score indicate its stronger overall performance in identifying both categories of mechanical properties accurately. The higher recall suggests that IRIN effectively captures more instances of coarse aggregate with bad mechanical properties, minimizing the risk of missing critical cases. The F1 score, which balances precision and recall, further highlights IRIN's ability to maintain reliable identification while ensuring a balance between false positives and false negatives.

In addition to the high classification performance, interpretability is a concern in an actual engineering problem. The IBRB automatically generated by IRIN shows the law for identifying the mechanical properties of coarse aggregate of reinforced recycled clay brick in concrete. One of the rules in IBRB is as follows.

Table 5. Performance index values of each classification method.

Index	Recall	Precision	F1_Score
MLP	60.00%	36.00%	45.00%
DTC	92.86%	94.05%	92.97%
SVC	35.71%	12.76%	18.80%
KNC	73.33%	76.43%	73.57%
RFC	92.86%	94.05%	92.97%
RIN	86.67%	86.67%	86.67%
IRIN	93.94%	93.93%	93.12%

R_1 : if the apparent density is{ (Low, 0.0), (Medium, [0.3213, 0.4912]), (High, [0.5087, 0.6786]) },
AND the absorption rate is{ (Weak, [0.5617, 1.0]), (Medium, [0.0, 0.4383]), (Strong, 0.0) },
AND the crushing value is{ (Low, [0.4864, 0.6831]), (Medium, [0.3169, 0.5136]), (High, 0.0) },
AND the substitution rate is{ (Low, 0.0), (Medium, [0.4378, 0.5515]), (High, [0.4484, 0.5621]) },
AND the glue-water ratio is{ (Low, 0.0), (Medium, [0.6141, 0.6255]), (High, [0.3745, 0.3859]) },
Then the mechanical property is{ (Good, [0.6211, 0.7289]), (Bad, [0.2196, 0.3789]) },
with rule weight is 0.8311
and attribute weight is{ (0.9062), (0.8523), (0.8912), (0.4536), (0.6611) }

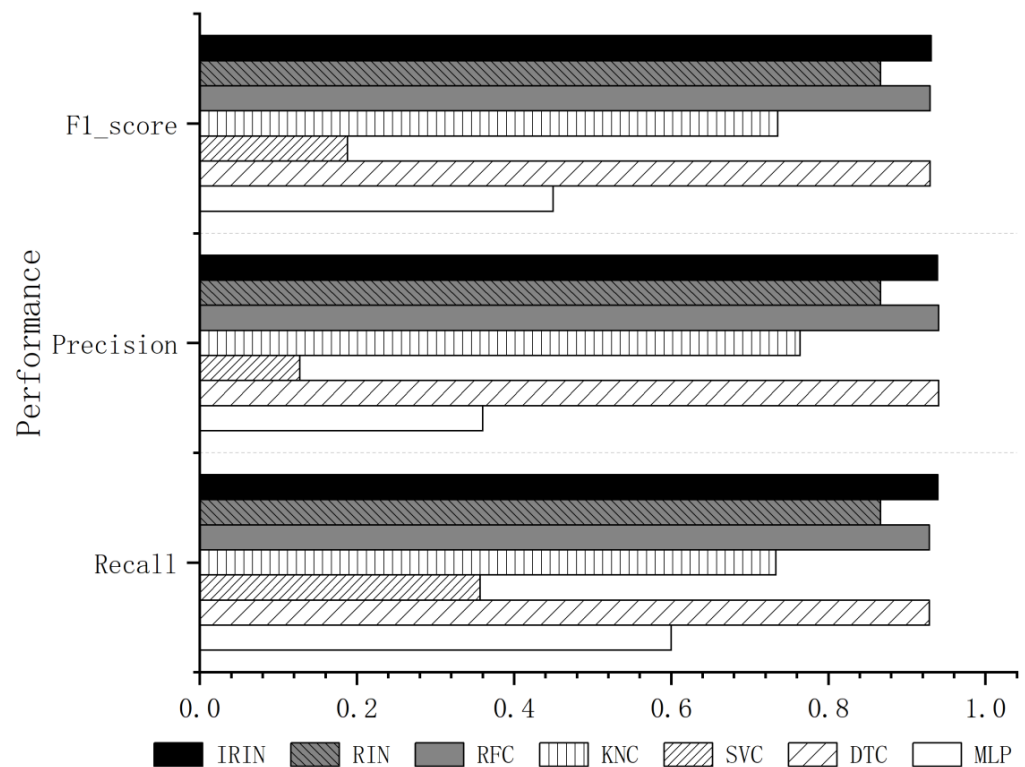


Figure 4. Comparison of performance indicators of various classification methods.

It is noted that when the expert or decision-maker believes that the mechanical properties of the rule should be more than 80% good, the expert or decision-maker can intervene in the rule to correct the result.

The experimental results show that the proposed IRIN method can effectively identify the mechanical properties of coarse aggregate of reinforced recycled clay brick in concrete and the IBRB can be intervened with to further improve the performance. Therefore, the proposed method can be applied to real-world scenarios.

7. Conclusions

In this paper, the IRIN model established by integrating the ideas of IBRB and neural network was proposed, to solve the problem of interval uncertainty and the inability of experts to accurately set a large number of parameters of IBRB. This model uses the inference mechanism of IBRB and updates parameters according to the gradient descent algorithm to automatically generate reasonable IBRB. This method not only effectively solves the complexity problem of parameter determination in a rule base, but also effectively handles interval uncertainty while maintaining interpretability.

Experimental results show that the proposed model has interpretability and intervention ability, improving the reliability of the model, and higher classification accuracy compared with other BRB systems. In addition, the IRIN system was used to identify the mechanical properties of reinforced recycled clay brick coarse aggregate in concrete, demonstrating that IRIN can effectively capture more instances of coarse aggregate with bad mechanical properties, minimizing the risk of missing critical cases.

However, since IRIN's learning algorithm only considers the boundaries of intervals, this method may lose part of the interval information, reducing the inference efficiency of IBRB and, thus, affecting the performance of IRIN. Therefore, in future research, the interval information of the interval belief structures will be considered to fully express the characteristics of the interval data.

Author Contributions: Methodology, Y.Z. (Yunxia Zhang) and Y.Z. (Yiming Zhong); formal analysis, X.W., Y.Z. (Yunxia Zhang), and J.B.; writing—original draft, Y.Z. (Yunxia Zhang) and Y.Z. (Yiming Zhong); writing—review and editing, X.W., Y.Z. (Yunxia Zhang), and Y.Z. (Yiming Zhong); data curation, Y.Z. (Yiming Zhong); validation, Y.Z. (Yunxia Zhang) and J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Hebei Provincial Natural Science Foundation (No. F2022210023, F2023210001), National Natural Science Foundation of China (NSFC) (No. 62341121), and the Projects of PhDs' Start-up Research of GDUP (No. 2023bsqd2005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhang, Y.; Tiño, P.; Leonardis, A.; Tang, K. A survey on neural network interpretability. *IEEE Trans. Emerg. Top. Comput. Intell.* **2021**, *5*, 726–742. [\[CrossRef\]](#)
2. Liu, Z.; Xu, F. Interpretable neural networks: Principles and applications. *Front. Artif. Intell.* **2023**, *6*, 974295. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Prokopowicz, P.; Mikołajewski, D. Fuzzy approach to computational classification of burnout—Preliminary findings. *Appl. Sci.* **2022**, *12*, 3767. [\[CrossRef\]](#)
4. Cao, J.; Zhou, T.; Zhi, S.; Lam, S.; Ren, G.; Zhang, Y.; Wang, Y.; Dong, Y.; Cai, J. Fuzzy inference system with interpretable fuzzy rules: Advancing explainable artificial intelligence for disease diagnosis—A comprehensive review. *Inf. Sci.* **2024**, *662*, 120212. [\[CrossRef\]](#)
5. Yang, R.; Zhao, Y.; Shi, Y. RPREC: A Radar Plot Recognition Algorithm Based on Adaptive Evidence Classification. *Appl. Sci.* **2023**, *13*, 12511. [\[CrossRef\]](#)
6. Xu, Z.; Lu, W.; Hu, Z.; Yan, W.; Xue, W.; Zhou, T.; Jiang, F. Decision-Refillable-Based Shared Feature-Guided Fuzzy Classification for Personal Thermal Comfort. *Appl. Sci.* **2023**, *13*, 6332. [\[CrossRef\]](#)
7. Gao, F.; Bi, C.; Bi, W.; Zhang, A. A new belief rule base inference methodology with interval information based on the interval evidential reasoning algorithm. *Appl. Intell.* **2023**, *53*, 12504–12520. [\[CrossRef\]](#)
8. Hu, G.; He, W.; Sun, C.; Zhu, H.; Li, K.; Jiang, L. Hierarchical belief rule-based model for imbalanced multi-classification. *Expert Syst. Appl.* **2023**, *216*, 119451. [\[CrossRef\]](#)
9. 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. Part Syst. Hum.* **2006**, *36*, 266–285. [\[CrossRef\]](#)
10. Liu, J.; Martinez, L.; Calzada, A.; Wang, H. A novel belief rule base representation, generation and its inference methodology. *Knowl.-Based Syst.* **2013**, *53*, 129–141. [\[CrossRef\]](#)
11. AbuDahab, K.; Xu, D.L.; Chen, Y.W. A new belief rule base knowledge representation scheme and inference methodology using the evidential reasoning rule for evidence combination. *Expert Syst. Appl.* **2016**, *51*, 218–230. [\[CrossRef\]](#)
12. Wang, Y.M.; Yang, J.B.; Xu, D.L.; Chin, K.S. The evidential reasoning approach for multiple attribute decision analysis using interval belief degrees. *Eur. J. Oper. Res.* **2006**, *175*, 35–66. [\[CrossRef\]](#)
13. Zhu, H.; Zhao, J.; Xu, Y.; Du, L. Interval-valued belief rule inference methodology based on evidential reasoning-IRIMER. *Int. J. Inf. Technol. Decis. Mak.* **2016**, *15*, 1345–1366. [\[CrossRef\]](#)
14. Liu, M.; He, W.; Zhou, G.; Zhu, H. A New Student Performance Prediction Method Based on Belief Rule Base with Automated Construction. *Mathematics* **2024**, *12*, 2418. [\[CrossRef\]](#)
15. Zhang, H.; Yang, R.; He, W.; Feng, Z. Cooperative performance assessment for multiagent systems based on the belief rule base with continuous inputs. *Inf. Sci.* **2024**, *676*, 120815. [\[CrossRef\]](#)
16. Fu, Y.G.; Lin, X.Y.; Fang, G.C.; Li, J.; Cai, H.Y.; Gong, X.T.; Wang, Y.M. A novel extended rule-based system based on K-Nearest Neighbor graph. *Inf. Sci.* **2024**, *662*, 120158. [\[CrossRef\]](#)
17. Huang, D.; Zhang, Y.; Lin, H.; Zou, L.; Liu, Z. Rule inference network model for classification. *ruan jian xue bao. J. Softw.* **2020**, *31*, 1063–1078.
18. Gao, F.; Bi, W. A fast belief rule base generation and reduction method for classification problems. *Int. J. Approx. Reason.* **2023**, *160*, 108964. [\[CrossRef\]](#)

19. Yang, L.H.; Ye, F.F.; Liu, J.; Wang, Y.M. Belief rule-base expert system with multilayer tree structure for complex problems modeling. *Expert Syst. Appl.* **2023**, *217*, 119567. [\[CrossRef\]](#)
20. Jiao, L.; Zhang, H.; Geng, X.; Pan, Q. Belief rule learning and reasoning for classification based on fuzzy belief decision tree. *Int. J. Approx. Reason.* **2024**, *175*, 109300. [\[CrossRef\]](#)
21. Fu, C.; Hou, B.; Xue, M.; Chang, L.; Liu, W. Extended belief rule-based system with accurate rule weights and efficient rule activation for diagnosis of thyroid nodules. *IEEE Trans. Syst. Man Cybern. Syst.* **2022**, *53*, 251–263. [\[CrossRef\]](#)
22. Ma, J.; Zhang, A.; Gao, F.; Bi, W.; Tang, C. A novel rule generation and activation method for extended belief rule-based system based on improved decision tree. *Appl. Intell.* **2023**, *53*, 7355–7368. [\[CrossRef\]](#)
23. Leung, Y.; Fischer, M.M.; Wu, W.Z.; Mi, J.S. A rough set approach for the discovery of classification rules in interval-valued information systems. *Int. J. Approx. Reason.* **2008**, *47*, 233–246. [\[CrossRef\]](#)
24. Zhang, Y.; Li, T.; Luo, C.; Zhang, J.; Chen, H. Incremental updating of rough approximations in interval-valued information systems under attribute generalization. *Inf. Sci.* **2016**, *373*, 461–475. [\[CrossRef\]](#)
25. Zhang, X.; Mei, C.; Chen, D.; Li, J. Multi-confidence rule acquisition and confidence-preserved attribute reduction in interval-valued decision systems. *Int. J. Approx. Reason.* **2014**, *55*, 1787–1804. [\[CrossRef\]](#)
26. Cheng, X.; Qian, G.; He, W.; Zhou, G. A Liquid Launch Vehicle Safety Assessment Model Based on Semi-Quantitative Interval Belief Rule Base. *Mathematics* **2022**, *10*, 4772. [\[CrossRef\]](#)
27. Han, P.; He, W.; Cao, Y.; Li, Y.; Mu, Q.; Wang, Y. Lithium-ion battery health assessment method based on belief rule base with interpretability. *Appl. Soft Comput.* **2023**, *138*, 110160. [\[CrossRef\]](#)
28. Si, Z.; Shen, J.; He, W. Lithium-Ion Battery Health Assessment Method Based on Double Optimization Belief Rule Base with Interpretability. *Batteries* **2024**, *10*, 323. [\[CrossRef\]](#)
29. Yang, L.H.; Liu, J.; Wang, Y.M.; Martínez, L. Extended belief-rule-based system with new activation rule determination and weight calculation for classification problems. *Appl. Soft Comput.* **2018**, *72*, 261–272. [\[CrossRef\]](#)
30. Gao, F.; Zhang, A.; Bi, W.; Ma, J. A greedy belief rule base generation and learning method for classification problem. *Appl. Soft Comput.* **2021**, *98*, 106856. [\[CrossRef\]](#)
31. Fu, Y.; Yin, Z.; Su, M.; Wu, Y.; Liu, G. Construction and reasoning approach of belief rule-base for classification base on decision tree. *IEEE Access* **2020**, *8*, 138046–138057. [\[CrossRef\]](#)
32. Yang, L.H.; Wang, Y.M.; Lan, Y.X.; Chen, L.; Fu, Y.G. A data envelopment analysis (DEA)-based method for rule reduction in extended belief-rule-based systems. *Knowl.-Based Syst.* **2017**, *123*, 174–187. [\[CrossRef\]](#)
33. Hao, G.; Feng, H.; Yu, H.; Zhao, Y.; Li, J.; Haotian, W. Calculation method of compressive strength of recycled concrete considering substitution rate and brick aggregate quality. *Hebei J. Ind. Sci. Technol.* **2023**, *40*, 202–209.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.