An Ensemble Extended Belief Rule Base Decision Model for Imbalanced Classification Problems

Yang, L-H., Ren, T-Y., Ye, F-F., Nicholl, P., Wang, Y-M., & Lu, H. (2022). An Ensemble Extended Belief Rule Base Decision Model for Imbalanced Classification Problems. *Knowledge-Based Systems*, *242*, 1-15. [108410]. https://doi.org/10.1016/j.knosys.2022.108410

Published in:
Knowledge-Based Systems

Publication Status:
Published (in print/issue): 22/04/2022

DOI:
https://doi.org/10.1016/j.knosys.2022.108410

Document Version
Author Accepted version

An Ensemble Extended Belief Rule Base Decision Model for Imbalanced Classification Problems

Long-Hao Yang [a], Tian-Yu Ren [a], Fei-Fei Ye [b, *], Peter Nicholl [c], Ying-Ming Wang [a], and Haitian Lu [d]

[a]School of Economics and Management, Fuzhou University, Fuzhou, 350116, China

[b]School of Public Management, Fujian Normal University, 350117, China

[c]School of Computing, Ulster University, Northern Ireland, United Kingdom

[d]School of Accounting and Finance, The Hong Kong Polytechnic University, Hong Kong, China

*The Corresponding Author.

Abstract: Class imbalance is a common problem in real-world applications and usually poses a major challenge to artificial intelligent (AI)-based decision models. The present work introduces a novel ensemble decision model which utilizes an explainable and fast-growing rule-based system, called extended belief rule base (EBRB) decision model, to alleviate the impact of class imbalance, where the proposed ensemble EBRB model includes two core components: a diversity-based base EBRB construction scheme and a consistency-based ensemble EBRB inference scheme. Specifically, for the purpose of enhancing diversity in the construction scheme, various kinds of oversampling techniques are applied to construct diverse base EBRBs firstly, followed by the calculation of attribute weights based on information gain. As for the inference scheme, the proposed ensemble EBRB model aims to produce inferential outputs not only integrating the rules activated from all base EBRBs, but also taking into consideration the consistency of the activated rules. In experimental study, twenty-six imbalanced classification datasets are used to demonstrate the effectiveness of the proposed ensemble EBRB decision model. Results demonstrate that the proposed model outperforms conventional EBRB systems and other typical imbalanced classifiers.

Keywords: Belief rule base; Imbalanced classification; Diversity; Inconsistency; Oversampling

1. Introduction

Aiming to label unseen data samples, classification tasks have become an important topic in machine learning community and play a fundamental role in modern industries and real-world applications. However, for the majority of machine learning algorithms, there still exist challenges for knowledge discovery from the datasets with skewed distribution, which are also known as imbalanced classification problems. In such problems, one or several classes (minority class) contain little data comparing to other classes (majority class), nevertheless, the samples of minority class are often the kernel samples and the machine learning algorithms fail to move their decision boundaries closer to the minority class since they were designed to minimize the misclassification cost on all training data. This is the reason why imbalanced classification problems poses a crucial challenge to artificial intelligent (AI)-based decision systems.

Imbalanced classification problems are prevalent in practical scenarios, such as defect detection [1], medical diagnosis [2], and fraud detection [3]. In the past decades, many attempts have been undertaken to address and solve imbalanced classification problems and these attempts can be categorized into: data-level, algorithm-level, and ensemble approaches. More details on the attempts can be found in Section 2. Note that the attempts of data-level and algorithm-level approaches are based on single classifiers, which may cause a great bias to the minority class and fail to generate a robust prediction result. The ensemble approach is becoming more popular and it combines the ensemble learning theorem with existing classifiers and oversampling techniques to provide accurate prediction results when handling imbalanced classification problems.

So far, many kinds of classifiers, such as support vector machine (SVM), artificial neural network (ANN), and *k*-nearest neighbor (KNN), have been applied to propose an ensemble classifier for dealing with imbalanced classification problems.

However, all these ensemble classifiers only focused on a desired accuracy for any given imbalanced dataset without taking into consideration the explainability of the process that a classifier produces a prediction output from input data. Recently, the extended belief rule base (EBRB) decision model [7] was used to propose a boosting-based ensemble EBRB classifier for imbalanced classification problems [43], where the EBRB decision model was designed as one integrated data and knowledge driven decision model with high interpretability and efficiency. In other words, the EBRB decision model is exactly a 'white-box' model and it can easily interpret the reasons why decision makers make such predictions.

However, the existing studies of EBRB decision model on imbalanced classification problems show that 1) few attempts have been done to develop a powerful classifier based on EBRB decision model for imbalanced classification problems; and 2) the existing ensemble EBRB classifier failed to consider high-efficiency and components diversity. This creates two challenges that must be investigated. The first challenge is how to ensure the components diversity of an ensemble EBRB decision model because the base EBRB (component) is directly generated from datasets so that it is hard to achieve the diversity of base EBRB. The second challenge is how to ensure the consistency of an ensemble EBRB decision model because the consistency among activated rules has notable influence on both minority and majority class accuracies. Considering the class overlap tends to be more serious in the context of class imbalance, ensuring consistency is still imperative for the ensemble EBRB decision model while distinguishing component outputs instead of maintaining consistency seems to be more feasible.

For the first challenge, the present work initiates a diversity-based base EBRB construction method. Since the diversity plays a fundamental role in improving the performance of an ensemble model, generating a diverse base EBRB is essential for the construction process of the ensemble EBRB decision model. However, common ensemble methods which were mainly designed for parametric models may not be sufficient enough to generate the diverse base EBRBs. To develop a novel method for constructing an ensemble EBRB decision model from an imbalanced dataset, in the present work, one random sampling technique is firstly utilized to extract several base datasets from original dataset. Various kinds of oversampling techniques are then used to rebalance those base datasets. When the base EBRBs are constructed using the base datasets, the information gain is used to calculate attribute weights for further improving the diversity of all base EBRBs.

For the second challenge, a consistency-based ensemble EBRB inference method is proposed to aggregate the rules activated from all base EBRBs. In the context of data imbalance, the class overlap must be a more serious problem when an ensemble classifier is constructed based on an imbalanced dataset and it has to cause the inconsistency between the activated rules. In other words, the immediate cause of hindering the performance of ensemble EBRB decision model is resulted from the class overlap. For this reason, a new definition regarding consistency is provided for the first time to effectively measure the consistency of the activated rules, and then the consistency is used together with the integrated belief degrees derived from all base EBRBs to produce final belief degrees. Owing to the proposed inference method, the ensemble EBRB decision model is able to handle imbalanced classification problems in a desired performance.

On the basis of the above-mentioned solutions for the two challenges, a novel ensemble EBRB decision model can be constructed by using the diversity-based base EBRB construction method to construct multiple base EBRBs from a given imbalanced dataset and the consistency-based ensemble EBRB inference method to produce an accurate output for any given query data. In the case study, four experiments with 26 imbalanced classification datasets are used to verity the effectiveness of the proposed ensemble EBRB decision model. Moreover, several ensemble EBRB decision models with different modeling setting and commonly used machine learning algorithms are applied to compare the performance of the proposed ensemble

EBRB decision model.

The novelties and contributions of this present study can be summarized below: 1) the use of multiple data sampling methods and information gain to improve the conventional EBRB construction scheme for imbalanced classification problems; 2) the consideration of the activated rules' consistency in each base EBRB to improve the conventional EBRB inference scheme for imbalanced classification problems; and 3) numerous preliminary conclusions are obtained to clarify the inherent features of the proposed ensemble EBRB decision model under imbalanced classification problems

The remainder of the present study is organized as follows: Section 2 introduces the related work of data imbalance. Section 3 gives an overview of the EBRB decision model and its challenges for imbalanced classification; Section 4 proposes a novel ensemble EBRB decision model for imbalanced classification; Section 5 provides an experimental study to validate the effectiveness of the proposed model. Finally, Section 6 draws conclusions from the study.

2. Related Works on Data Imbalance

To date, the widely used approaches for imbalanced classification can be segmented into data-level approaches, algorithm -level approaches, and ensemble approaches. The detailed reviews of these three categories are provided as follows:

(1) Data-level approaches. Data-level approaches are mainly based on generating minority class samples or removing majority class samples to rebalance data distribution. These two kinds of data processing ways are called oversampling and undersampling techniques. Among existing oversampling techniques, synthetic minority oversampling technique (SMOTE) [8] is a well-established over-sampling technique and many kinds of variants have been developed for effectively rebalance data distribution. For example, Han *et al*. [6] proposed the borderline-SMOTE, which only oversamples the minority class samples near the borderline. Experiments showed that the borderline-SMOTE could achieve better performance than SMOTE and random oversampling technique. Afterwards, Bunkhumpornpat *et al*. [5] proposed the Safe-Level-SMOTE based on the idea that the minority class samples along the same line with different safe levels are oversampled. They demonstrated that the Safe-Level-SMOTE could have better precision and F-value than SMOTE and Borderline-SMOTE when a decision tree is applied as a classifier. He *et al*. [4] presented the adaptive synthetic (ADASYN) oversampling technique for learning from an imbalanced dataset, where the ADASYN not only can generate synthetic data for the minority class, but also is able to shift the classifier decision boundary to be more focused on those difficult to learn data samples. Recently, Juez-Gil *et al*. [44] proposed the Approx-SMOTE which is a parallel implementation of the SMOTE under the Apache Spark framework. The big data related experiments demonstrated that the Approx-SMOTE is able to achieve up to 30 times faster execution times without sacrificing the improved classification performance offered by the original SMOTE. Comparing to oversampling techniques, few attentions have been paid to undersampling techniques because it is accompanied by information loss owing to removing majority class samples. The representative studies have: Anand *et al*. [48] proposed an undersampling technique targeting the boundary data samples which are always challenging to deal with for any classifiers. Lin *et al*. [27] proposed a clustering-based undersampling technique based on the process that only majority class samples is divided into several clusters and the cluster centers or their nearest neighbors is utilized to represent the majority class.

(2) Algorithm-level approaches. Algorithm-level approaches are mainly based on the modification of some standard algorithms to address imbalanced classification problems. This kind of approaches can be categorized into cost-sensitive learning (CSL) and algorithmic classifiers modifications. The representative studies of the former includes: Yu *et al*. [28] developed a support vector machine (SVM) decision threshold adjustment algorithm called SVM-OTHR. This method first

generated several new hyperplane positions for the misclassified minority class samples and then exhaustively searched for the best position; Zhang *et al*. [29] combined the oversampling techniques with CSL to address imbalanced datasets, where the higher-cost minority samples is replicated according to the given cost matrices to balance the distribution of minority class. Due to that the large margin distribution machine (LDM) failed to handle imbalanced datasets, a cost-sensitive LDM (CS-LDM) was proposed in [30] to improve the detection rate of minority class. The main idea is to increase the margin weight of the minority class in the margin mean and the misclassification penalty of the minority class to let the minority class has a larger margin to the separator. For the algorithmic classifiers' modifications, Zhang *et al*. [31] proposed a scaling kernel-based SVM to deal with multi-class imbalanced data classification problems. Specifically, the proposed method first used SVM to gain an approximate hyperplane. Then, a scaling kernel function was presented and the parameters are calculated using the chi-square test and weighting factors. Kim *et al*. [32] proposed a modified weight-k-nearest neighbor classifier called α-wkNN, where the α-wkNN should learn an optimized α value to be decision threshold from imbalanced datasets. Later, a novel support vector data description (SVDD) model [33] was proposed by combining binary tree (BT) to handle imbalanced multi-classification problems.

(3) Ensemble approaches. Ensemble approaches are mainly based on ensemble learning theorem to handle imbalanced datasets and have gained popularity in recent years. Since the ensemble approaches aim to provide a framework of constructing classifiers for imbalanced datasets, the algorithm-level and/or data-level approaches should be used together [47]. For the algorithm-level-based ensemble approaches, Sun *et al*. [34] developed three different cost items to update the weight of AdaBoost for imbalanced datasets. Experimental results showed that the proposed ensemble classifiers increased more weights on misclassified minority samples and less on majority samples; Ali *et al*. [35] presented a cost-sensitive ensemble model called Can-CSC-GBE for a breast cancer dataset by incorporating CSL with GentleBoost, AdaBoostM1, and Bagging. They demonstrated that the cost of misclassifying a cancer patient as non-cancer equals to the imbalance ratio. Afterwards, a more advanced cost-sensitive ensemble classifier tried to learn from imbalanced datasets without a prior cost was proposed in [36]. In this study, the imbalanced classification problem was reformulated into a partial ranking problem that learned a non-parametric scoring function to maximize the difference between the majority and minority classes. Comparing to algorithm-level-based ensemble approaches, data-level-based ensemble approaches are more common and general in the previous studies. For example, Barandela *et al*. [37] proposed UnderBagging based on the way of randomly undersampling the majority class samples in each iteration of the Bagging algorithm while all minority class samples are kept. Wang and Yao [39] used SMOTE in each iteration of the Bagging algorithm to propose SMOTEBagging, where the size of oversampled dataset is two times that of original majority class. The first half is the bootstrapped replica of majority examples, whereas the second half is obtained by SMOTE or random oversampling depending on the oversampling rate; Later, Chawla *et al*. [40] utilized SMOTE to generate synthetic minority class data for proposing a new data-level-based ensemble approach called SMOTEBoost. The other representative data-level-based ensemble approaches include EasyEnsembles [38], ECO-Ensemble [41] and DPHS-MDS [42]. Recently, the EBRB decision model, which has been widely used to handle classification problems and demonstrated its explainability, efficiency, and accuracy over other traditional classifiers [10-23], was introduced to propose a boosting-based ensemble EBRB classifier for dealing with imbalanced datasets [43]. Experiments showed that the proposed EBRB classifier significantly improves F-value compared to other imbalanced classification algorithms.

It can be found from the above-mentioned related works on data imbalance that the ensemble approaches are more advanced and popular approaches. As the latest and representative study of the ensemble approaches, the boosting-based ensemble EBRB decision model showed its potential over other ensemble classifiers. However, this kind of EBRB decision model is a low-efficiency classifier because it must include an iterative training process to determine the values of parameters. Besides, it failed to distinguish the reliability of each component of the boosting-based EBRB classifier. Hence, in order to make full use of EBRB decision model, it is desirable to develop a novel ensemble classifier based on EBRB decision model for handling imbalanced classification problems.

3. EBRB Decision Model for Classification Problems: Overview and Challenges

In this section, an overview of the EBRB decision model is outlined. After that, the challenges are summarized for the EBRB decision model for handling imbalanced classification problems.

3.1. Basics of Extended Belief Rule Base (EBRB)

As the rule base of an EBRB decision model, EBRB contains M antecedent attributes U_i (i=1,…, M) with J_i reference values $A_{i,j}$ (j=1,…, J_i) and one consequent attribute D having N consequents D_n (n=1,…, N). Consequently, the kth (k=1,…, L) extended belief rule in EBRB can be represented as:

$$\begin{gathered} R_k : IF\ U_1\ is\ \{(A_{1,j}, \alpha_{1,j}^k); j=1,\ldots,J_1\} \wedge \cdots \wedge U_M\ is\ \{(A_{M,j}, \alpha_{M,j}^k); j=1,\ldots,J_M\} \\ THEN\ D\ is\ \{(D_n, \beta_n^k); n=1,\ldots,N\}, with\ \theta_k\ and\ \{\delta_1,\ldots,\delta_M\} \end{gathered} \tag{1}$$

where $\alpha_{i,j}^k (0 \le \alpha_{i,j}^k \le 1)$ and $\beta_n^k (0 \le \beta_n^k \le 1)$ denote the belief degree of reference value $A_{i,j}$ and consequent D_n in the kth rule, respectively. In addition, θ_k denotes the weight of the kth rule and δ_i denotes the weight of the ith antecedent attribute.

From Eq. (1), it is worth noting that the extended belief rule is a flexible and advanced rule representation scheme because of the belief structure, which reflects that the kth rule is considered to be incomplete when $\sum_{j=1}^{J_i} \alpha_{i,j}^k < 1$ or $\sum_{n=1}^{N} \beta_n^k < 1$, and it is considered to be complete when $\sum_{j=1}^{J_i} \alpha_{i,j}^k = 1$ and $\sum_{n=1}^{N} \beta_n^k = 1$.

In an EBRB decision model, the construction of an EBRB for classification problems includes the three steps below:

Step 1: To initialize basic parameters. According to prior knowledge, the value of the basic parameters, including M attribute weights δ_i (i=1,…, M), $\sum_{i=1}^{M} J_i$ utility values $u(A_{i,j})$ (j=1,…, J_i) related to M antecedent attributes, and N classes related to N consequents D_n (n=1,…, N) in the consequent attribute.

Step 2: To generate belief distributions. Suppose that there are L historical input-output data pairs <$\boldsymbol{x_k}$, y_k> (k=1,…, L), where $\boldsymbol{x_k}$=($x_{k,1}$,…, $x_{k,M}$) denotes the kth input vector and $x_{k,i}$ is the kth input data on the ith antecedent attribute, y_k denotes the kth output data on the consequent attribute. By utilizing the utility-based transformation technique[39], these L historical data pairs are converted into belief distributions, in which the belief distribution of the ith antecedent attribute is as follows:

$$S(x_{k,i}) = \{(A_{i,j} \alpha_{i,j}^k), j=1,\ldots,J_i\} \tag{2}$$

where

$$\alpha_{i,j}^k = \frac{u(A_{i,j+1}) - x_{k,i}}{u(A_{i,j+1}) - u(A_{i,j})}\ and\ \alpha_{i,j+1}^k = 1 - \alpha_{i,j}^k, if\ u(A_{i,j}) \le x_{k,i} \le u(A_{i,j+1}) \tag{3}$$

$$\alpha_{i,t}^k = 0\ for\ t=1,\ldots,J_i\ and\ t \ne j, j+1 \tag{4}$$

The belief distribution of the consequent attribute can be obtained by:

$$S(y_k) = \{(D_n, \beta_n^k); n=1,\ldots,N\} \tag{5}$$

$$\beta_n^k = \begin{cases} 1, if\ y_k = D_n \\ 0, otherwise \end{cases} \tag{6}$$

Step 3: To calculate rule weights. After obtaining $L\times(M+1)$ belief distributions from the L historical input-output data pairs, all these belief distributions can form L extended belief rules, where the weight of the kth (k=1,…, L) extended belief rule is calculated by

$$\theta_k = 1 - \frac{ID(R_k)}{\sum_{j=1}^{L} ID(R_j)} \tag{7}$$

where $Incons(R_k)$ denotes the inconsistency degree between the kth rule to others and it is calculated by

$$Incons(R_k) = \sum_{l=1,l\neq k}^{L} (1-\exp\{(\frac{SRA(R_l,R_k)}{SRC(R_l,R_k)}-1)^2 \left(SRA(R_l,R_k)\right)^2\}) \tag{8}$$

where $SRA(R_l, R_k)$ and $SRC(R_l, R_k)$ denote the similarity of rule similarity (SRA) and rule consequent (SRC), and they are calculated by:

$$SRA(R_l,R_k) = 1-\max_{i=1\ldots M}\{\sqrt{\sum_{j=1}^{J_i}(\alpha_{i,j}^l-\alpha_{i,j}^k)^2}\}; l=1,\ldots,L; l\neq k \tag{9}$$

$$SRC(R_l,R_k) = 1-\sqrt{\sum_{j=1}^{J_l}(\beta_n^l-\beta_n^k)^2}; l=1,\ldots,L; l\neq k \tag{10}$$

3.2. Inference scheme for EBRB decision model

After the construction of EBRB, the EBRB decision model can produce an inferential class to reply any given input data based on the rule-based inference procedure below:

Step 1: To calculate individual matching degrees. Suppose that the new input vector is $\boldsymbol{x}=(x_1,\ldots, x_M)$, each input data x_i (i=1,…, M) in $\boldsymbol{x}$ should be transformed into the following belief distribution based on Eq. (3) and Eq. (4):

$$S(x_i) = \{(A_{i,j},\alpha_{i,j}); j=1,\ldots J_i\} \tag{11}$$

Then, the individual matching degree of the kth rule on the ith antecedent attribute can be calculated by:

$$S^k(x_i,U_i) = \begin{cases} \frac{\lambda-d_i^k}{\lambda}, d_i^k \leq \lambda \\ 0 \quad\quad , d_i^k > \lambda \end{cases}, d_i^k = \sqrt{\sum_{j=1}^{J_i}(\alpha_{i,j}-\alpha_{i,j}^k)^2}, \lambda\in[0,\sqrt{2}] \tag{12}$$

where λ denotes the activation factor [17]; d_i^k denotes Euclidean distance between two belief distributions. Note that, to ensure the range of individual matching degree being in [0, 1], Eq. (12) includes an additional normalization compared to the calculation of individual matching degrees detailed in [17].

Step 2: To calculate activation weights. Based on rule weights θ_k, attribute weights δ_i, and individual matching degree $S^k(x_i, U_i)$, the activation weight of the kth rule, denoted as w_k, can be calculated by:

$$w_k = \frac{\theta_k\prod_{i=1}^{M}S^k(x_i,U_i)^{\bar{\delta}_i}}{\sum_{l=1}^{L}\theta_l\prod_{i=1}^{M}S^l(x_i,U_i)^{\bar{\delta}_i}}, \bar{\delta}_i = \frac{\delta_i}{\max_{j=1,\ldots,M}\{\delta_j\}} \tag{13}$$

Step 3: To integrate activated rules. The distributed inferential output to reply the given input data $\boldsymbol{x}$ can be obtained by integrating all activated rules, whose activation weights are greater than 0, using the Evidence Reasoning for Classification (ER-C) algorithm [40] below:

$$\beta_n = \prod_{k=1}^{L}(w_k\beta_n^k + 1 - w_k\sum_{i=1}^{N}\beta_i^k) \tag{14}$$

The inferential class of the input data $\boldsymbol{x}$ can be calculated by:

$$f(\boldsymbol{x}) = D_n, n = \arg\max_{t=1,...,N}\{\beta_t\} \tag{15}$$

3.3. Challenges of EBRB decision model for imbalanced classification

One of reasons why a classifier rarely performs well on imbalanced classification problems is that the boundary of the classifier has to move closer to minority class due to the sightless guidance of obtaining higher accuracy on entire data. This issue can be also found in the EBRB decision model, thus the following challenges should be discussed and addressed when an ensemble approach is used to improve the EBRB decision model for handling imbalanced classification problems.

***Challenge* 1**: The diversity of base EBRBs under imbalanced classification problems.

Data sampling-based ensemble classifiers have demonstrated the great potential in handling imbalanced classification problems. As mentioned in [9], the diversity of base classifiers plays a fundamental role in improving the performance of ensemble classifiers. However, the diversity of the EBRB-based ensemble classifier is typically challenging to be achieved under the framework of data sampling. This is because the construction of EBRB is dependent on training data, namely the diversity of training data has a significant influence on the diversity of EBRB, leading to the dilemma that the diversity loss of the base EBRB has to be aggravated due to the unbalance of datasets. Hence, in the present work, the first challenge is to propose an effective data sampling method for constructing diverse base EBRBs, so all these base EBRBs can be integrated to develop an accurate ensemble EBRB decision model for imbalanced classification problems.

***Challenge* 2**: The consistency of ensemble EBRB inference under imbalanced classification problems.

The inference core of the EBRB decision model is to activate the similar rules comparing to the given input and integrate these activated rules for producing an inferential output. Thus, the consistency of activated rules has notable influence on the accuracy of the EBRB decision model. Owing to the fact that an ensemble EBRB decision model consists of multiple base EBRBs, keeping consistency of the activated rules in the ensemble EBRB decision model become more complex because the consistency of the activated rules obtained from all base EBRBs has to be considered simultaneously. In addition, the overlap region of data can easily weaken the performance of the EBRB decision model, especially for the imbalanced classification problems, because the data with majority class are usually the dominate data in overlap region, which makes that the rules generated from the data with minority class are regarded as disturbance to decrease the consistency of the activated rules, finally resulting in the low accuracy of the data with minority class.

For the above-mentioned challenges, it is clear from the related works shown in Section 2 that existing studies still cannot properly construct an ensemble EBRB decision model for imbalanced classification problems, because all these attempts fail to meet the requirement of the two challenges. Thus, in this study, a new ensemble EBRB decision model for imbalanced classification problems is developed to overcome the two challenges.

4. Ensemble EBRB Decision Model for Imbalanced Classification Problems

In order to overcome the two challenges discussed in Section 3.3, a novel ensemble classifier is proposed based on the EBRB decision model for handling imbalanced classification problems in Sections 4.1 to 4.3.

4.1. Framework of ensemble EBRB decision model

In this section, an ensemble EBRB decision model is proposed for imbalanced classification problems and it consists of two core components: diversity-based base EBRB construction and consistency-based ensemble EBRB inference. Fig. 1 shows the basic framework of the ensemble EBRB decision model.

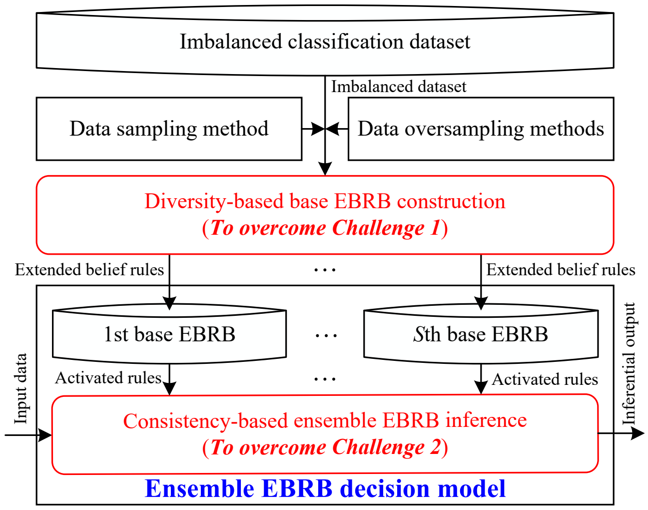

Fig. 1. Framework of ensemble EBRB decision model for imbalanced classification

From Fig. 1, the steps of the ensemble EBRB decision model for imbalanced classification are provided as follows:

***Step* 1**: Diversity-based base EBRB construction. For handling imbalanced classification problems, diverse data sampling methods are necessary and important approaches to construct base classifiers in the aim of enhancing the diversity among base classifiers and eliminating the unbalance of datasets. Hence, the random data sampling and oversampling methods detailed in Section 4.2 are used together to improve the traditional EBRB construction process. Moreover, the attribute weights of each base EBRB are calculated based on information gain to further enhance the diversity of base EBRBs.

***Step* 2**: Consistency-based ensemble EBRB inference. After constructing multiple base EBRBs based on Step 1, all of these base EBRBs should be used together to classify any given input data. As each base EBRB is an independent unit in the ensemble EBRB decision model, the consistency of the activated rules in each base EBRB should be measured independently. Afterwards, the consistency-based integrated belief degree is defined to take into consideration the consistency of all base EBRBs for producing an accurate inferential output. The details of this step can be found in Section 4.3.

4.2. Diversity-based base EBRB construction

As mentioned, it is meaningful and imperative to develop an ensemble EBRB decision model from an imbalanced dataset. Considering data sampling-based ensemble classifiers were widely used in previous studies with high efficiency for imbalanced classification problems, in this section, a random data sampling method together with different kinds of data oversampling methods are introduced to construct diverse base EBRBs based on traditional EBRB construction process shown in Section 3.1. Afterwards, in order to further enhance the diversity of the base EBRBs, all these base EBRBs are assigned to have different attribute weights based on the information gain which indicates how much information an antecedent attribute reflects about the final output. Such that, the attribute weight is calculated by subtracting the entropy of a particular attribute inside the dataset from the entropy of the whole dataset, where the entropy is the degree of disorder or randomness indicating how impure or uncertain the data in the set is. The process of diversity-based base EBRB construction is shown in Fig. 2.

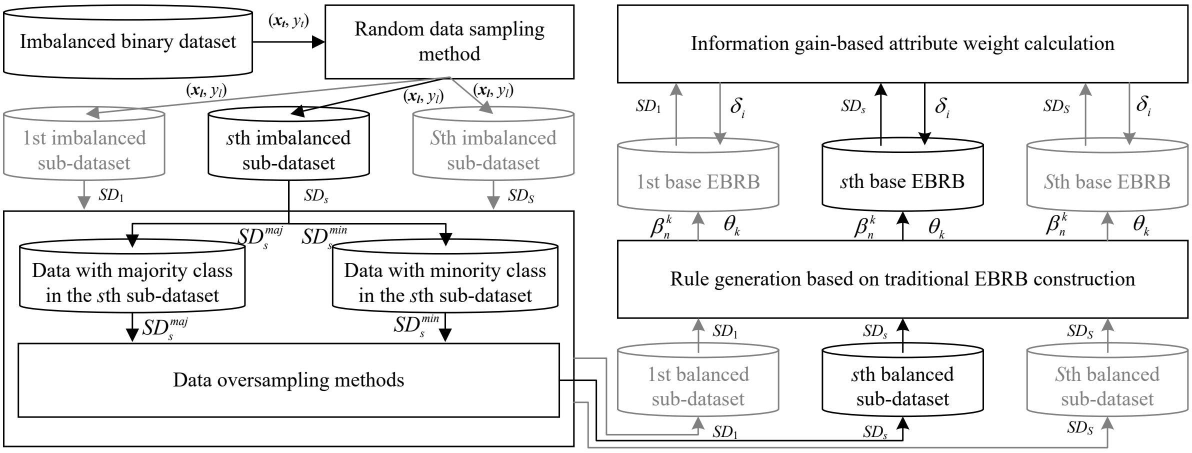

Fig. 2. The proposed diversity-based base EBRB construction process

According to Fig. 2, the base EBRB can be constructed by the following steps:

***Step* 1**: To generate S sub-datasets from imbalanced classification dataset. Suppose that an imbalanced dataset has T data with M inputs U_i (i=1,…, M) and one output D, denoted as ($\boldsymbol{x_t}$, y_t) and $\boldsymbol{x_t}$= ($x_{t,1}$,…, $x_{t,M}$). A total of S sub-datasets, denoted as SD_s (s=1,…, S), can be generated according to the following pseudocode and the size of SD_s is assumed as L:

Pseudocode 1: generation of S base datasets from original dataset

For s = 1 to S do
  Initialize the sth base dataset SD_s ={}
  For k = 1 to T do
    Select a data ($\boldsymbol{x_l}$, y_l) from {($\boldsymbol{x_t}$, y_t); t=1, …, T} by random way
    If ($\boldsymbol{x_l}$, y_l) is not in SD_s do
      $SD_s = SD_s \cup \{(\boldsymbol{x_l}, y_l)\}$
    End if
  End for
End for

***Step* 2**: To balance the number of data for minority class based on data oversampling methods. For the sth sub-dataset SD_s obtained from *Step* 1, suppose that the number of data of minority class and majority class is L_{min} and L_{maj}, respectively. Thus, the total number of new data generated by one data oversampling method for minority class is L_{new}=L_{maj}-L_{min}. Here, it is worth noting that different kinds of data oversampling methods should be used to balance the data distribution of SD_s, so the oversampled SD_s has enough diversity to ensure the performance of ensemble EBRB decision model.

***Step* 3**: To calculate belief distributions and rule weights. For the sth sub-dataset SD_s obtained from *Step* 2, each input and output data pair ($\boldsymbol{x_k}$, y_k) (($\boldsymbol{x_k}$, y_k)∈SD_s) is used to generate an extended belief rule for the sth base EBRB by the following pseudocode:

Pseudocode 2: generation of belief distributions and rule weights

For each ($\boldsymbol{x_k}$, y_k) in SD_s do
  For each $x_{k,i}$ in $\boldsymbol{x_k}$ do
    Calculate $S(x_{k,i}) = \{(A_{i,j} \alpha_{i,j}^k), j = 1,..., J_i\}$ based on Eqs. (3) and (4)
  End for
  Calculate $S(y_k) = \{(D_n, \beta_n^k); n = 1,..., N\}$ based on Eq. (6).
  Calculate θ_k from $S(x_{k,i})$ and $S(y_k)$ based on Eqs. (7) to (10)

Form the kth extended belief rule R_k using θ_k, $S(x_{k,i})$, and $S(y_k)$

End for

***Step* 4**: To calculate attribute weights. For the sth base EBRB obtained from *Step* 3, the weight of each attribute is calculated using the sth sub-dataset SD_s obtained from *Step* 2 and the following steps:

***Step* 4.1**: To calculate information entropy. The information entropy of BD_s is calculated by

$$E(SD_s) = p_{min} \log_2 p_{min} + p_{maj} \log_2 p_{maj} \tag{16}$$

where p_{min} and p_{maj} denote the proportion of minority and majority class data, and they are calculated by

$$p_{min} = \frac{L_{min}}{L_{min} + L_{maj}}, p_{maj} = \frac{L_{maj}}{L_{min} + L_{maj}} \tag{17}$$

***Step* 4.2**: To calculate information gain. Suppose that there are J_i reference values $A_{i,j}$ (j=1,…, J_i) for the ith (i=1,…, M) antecedent attribute U_i. The information gain of U_i is calculated by

$$G(SD_s, U_i) = E(SD_s) - \sum_{j=1}^{J_i} \frac{\left|SD_s^{i,j}\right|}{\left|SD_s\right|} E(SD_s^{i,j}) \tag{18}$$

where $SD_s^{i,j}$ denotes the subset of the input-output data pair ($\boldsymbol{x_k}$, y_k) (($\boldsymbol{x_k}$, y_k)$\in SD_s$) whose corresponding belief distribution $\{(A_{i,j}\alpha_{i,j}^k), j=1,...,J_i\}$ satisfying $\alpha_{i,j}^k \geq \alpha_{i,t}^k$ (t=1, …, J_i), namely

$$SD_s^{i,j} = \{(\boldsymbol{x_k}, y_k) \in SD_s \mid \alpha_{i,j}^k \geq \alpha_{i,t}^k; t = 1,...,J_i\} \tag{19}$$

***Step* 4.3**: To calculate attribute weights. The weight δ_i of the ith antecedent attribute is calculated by

$$\delta_i = \frac{G(SD_s, U_i)}{\max_{t=1,...,M}\{G(SD_s, U_t)\}} \tag{20}$$

In order to show the diversity of the base EBRBs, a benchmark binary imbalanced dataset, named new-thyroid-1 [24] containing 35 positive and 180 negative samples with 5 features T3resin, Thyroxin, Triiodothyronine, Thyroid stimulating, and TSH value, is used to perform a case study with three kinds of data oversampling methods, including ADASYN [6], Borderline-SMOTE[5], and Safe-Level-SMOTE[4]. The Kohavi-Wolpert variance (KWV)[45, 46] is used to evaluate the difference of the inferential outputs of all base EBRBs in the ensemble EBRB decision model within the range [0, 0.25], and 0 denotes the lowest diversity of the ensemble model because all base EBRBs have the same inferential output; 0.25 denotes the highest deversity of the ensemble model when half of all base EBRBs have the same inferential output. The results are shown in Table 1 when considering different options to generate 30 base EBRBs for an ensemble EBRB decision model.

Table 1. Descriptions of ensemble EBRB decision model with different construction options

Ensemble EBRB decision model	Step 1	Step 2			Step 3	Step 4	KWV
		ADASYN	Borderline-SMOTE	Safe-Level-SMOTE			
Option 1	√	×	×	×	√	×	0.0044
Option 2	√	×	×	×	√	√	0.0047
Option 3	√	×	√	×	√	×	0.0080
Option 4	√	√	×	×	√	×	0.0173
Option 5	√	×	×	√	√	×	0.0196
Option 6	√	√	√	√	√	√	0.0438

As shown in Table 1, the ensemble EBRB decision model constructed by option 1 has the lowest KWV, which means that the predictions based on different base EBRBs are almost the same to each other. When the attribute weight calculation method detailed in Step 4 is considered, the corresponding KWV increases. Moreover, by considering Safe-Level SMOTE[5], Borderline-SMOTE [6], and ADASYN [4] to balance the imbalanced base datasets, the KWV continues to increase comparing to options 1 and 2. For the option 6, which includes the entire steps, the KWV is much greater than other options, even ten times of option 1. Therefore, it is clear that the proposed diversity-based base EBRB construction can effectively enhance the diversity of the ensemble EBRB decision model.

4.3. Consistency-based ensemble EBRB inference

From the previous studies on EBRB decision model [25], the consistency among activated rules is a crucial influence to weaken the classification accuracy of an EBRB decision model. Hence, the consistency is taken into consideration in the inference result of EBRB decision model. Specifically, the inference result obtained from the sth base EBRB is regarded as an unreliable result when the set of activated rules from the sth base EBRB has a low consistency. From this viewpoint, a new definition for the consistency of activated rules is given as follows:

***Definition* 1** (Consistency of activated rules). Suppose that AR_s is the set of activated rules for the sth base EBRB, D_n is the nth (n=1, ..., N) consequent of the consequent attribute and $\beta_{n,k}$ is the belief degree of the kth activated rule on the nth consequent. The consistency of activated rules AR_s is defined as:

$$C(AR_s) = \frac{\max_{n=1,\dots,N}\{C_n\}}{|AR_s|} \tag{24}$$

where C_n is given by:

$$C_n = \left| D_n ; n = \arg(\max_{i=1,\dots,N}\{\beta_{i,k}\}) \right| \tag{25}$$

Taking a binary classification problem with two classes $\{D_1, D_2\}$ for example, suppose the number of activated rules AR_s having the maximum belief degree on D_1 and D_2 is 8 and 2, namely $|AR_s|$=10, C_1=8 and C_2=2, the consistency of activate rules AR_s is therefore calculated by $C(AR_s)$=8/10=0.8.

***Definition* 2** (Consistency-based integrated belief degree). Based on Definition 1, when the integrated belief degree of the sth base EBRB to reply the input data $\boldsymbol{x}$ is assumed to be β_n^s (n=1,…, N) obtained from Eq. (14), the consistency- based integrated belief degree of S base EBRBs can be calculated by:

$$\beta_n = \frac{\sum_{s=1}^{S} C(AR_s)\beta_n^s}{\sum_{s=1}^{S} C(AR_s)} \tag{26}$$

On the basis of Definition 1 and Definition 2, the specific procedure of consistency-based ensemble EBRB inference is shown in Fig. 4 and the following pseudocode when S base EBRBs Φ_s (s=1,…, S) obtained from Section 4.1 are used to reply the given input data $\boldsymbol{x}$.

Pseudocode 3: **Consistency-based ensemble EBRB inference**

```
For each base EBRB Φs in {Φs; s=1,…, S} do
    ARs={}
    For each rule Rk in base EBRBΦs do
        Calculate activation weight wk based on Eqs. (11) to (13).
```

If $w_k > 0$ then

$AR_s = AR_s \cup R_k$

End if

End for

Calculate the consistency $C(AR_s)$ based on AR_s and Definition 1

Calculate integrated belief degrees $\{(D_n, \beta_n^s); n=1, \ldots, N\}$ based on AR_s and Eq. (14)

End for

Calculate consistency-based integrated belief degrees β_n based on $C(AR_s)$, β_n^s, and Definition 2.

Produce an inference output $f(\boldsymbol{x})$ to reply input data $\boldsymbol{x}$ based on Eq. (15)

For the earlier mentioned ensemble EBRB inference method, it is worth noting that the class overlap problem tends to be more serious in the context of data imbalance and this is the direct reason of decreasing the accuracy of EBRB decision model. By measuring the consistency of the activated rules derived from each base EBRB, it is helpful to distinguish the unreliable prediction outputs caused by inconsistent activated rules. Furthermore, by integrating these consistencies in the ensemble EBRB decision model, it is helpful to improve the accuracy in both majority class and minority class.

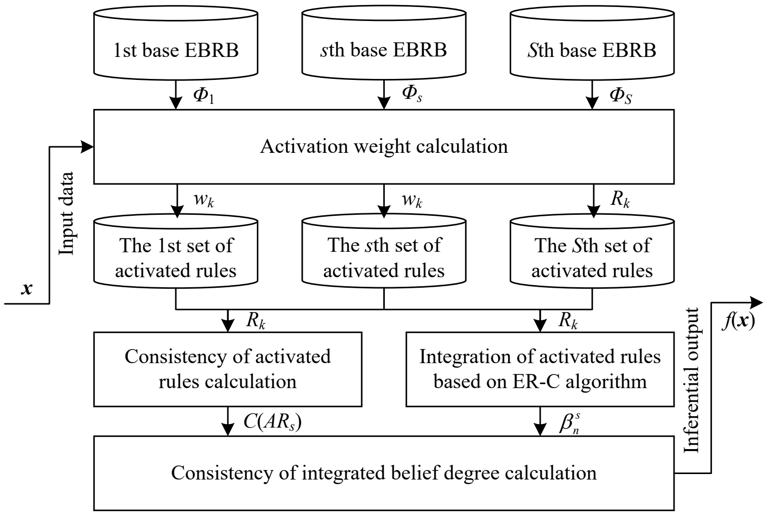

Fig.4. The proposed consistency-based ensemble EBRB inference process

5. Experimental study

In this section, an experimental study is performed according to the following aspects: Section 5.1 shows datasets and experimental settings; Section 5.2 compares ensemble and traditional EBRB decision models; Sections 5.3 and 5.4 analyze the function of data oversampling and attribute weight calculation on the ensemble EBRB decision model; and Section 5.5 provides comparative analysis for conventional classifiers.

5.1. Datasets and experimental settings

Twenty-six imbalanced classification datasets are utilized as benchmark datasets to evaluate the performance of the proposed ensemble EBRB decision model and all these datasets are obtained from the KEEL repository [24]. Table 2 provides the main characteristics of these 26 datasets, including the number of data (#Data) within the range [215, 1484], the number of attributes (#Attr) within the range [5, 18], and the imbalance ratio (#IR) within the range [1.86, 41.4].

Table 2. 26 imbalanced classification datasets

No	Datasets	Abbreviation	#Data	#Attr	#IR	No	Datasets	Abbreviation	#Data	#Attr	#IR
1	Ecoli1	Eco-1	336	7	1.86	14	Yeast4	Yea-3	1484	8	28.1
2	Ecoli2	Eco-2	336	7	5.46	15	Yeast6	Yea-4	1484	8	41.4
3	Ecoli3	Eco-3	336	7	8.60	16	Yeast1vs7	Yea-5	459	7	14.3
4	Ecoli4	Eco-4	336	7	15.80	17	Yeast2vs4	Yea-6	514	8	9.08
5	Ecoli0146vs5	Eco-5	280	6	13	18	Yeast2vs8	Yea-7	482	8	23.1
6	Ecoli01vs235	Eco-6	244	7	9.17	19	Yeast0359vs78	Yea-8	506	8	9.12
7	Ecoli01vs5	Eco-7	240	6	11	20	Yeast0256vs3789	Yea-9	1004	8	9.14
8	Ecoli0147vs56	Eco-8	332	6	12.28	21	Yeast02579vs368	Yea-10	1004	8	9.14
9	Ecoli0234vs5	Eco-9	202	7	9.1	22	Vehicle1	Veh-1	846	18	2.9
10	Ecoli0267vs35	Eco-10	224	7	9.18	23	Vehicle3	Veh-2	846	18	2.99
11	Ecoli0347vs56	Eco-11	257	7	9.28	24	New-thyroid-1	New-1	215	5	5.14
12	Yeast1	Yea-1	1484	8	2.46	25	New-thyroid-2	New-2	215	5	5.14
13	Yeast3	Yea-2	1484	8	28.1	26	Pima	Pim	768	8	1.87

To develop comparison in multiple aspects and diminish randomness, 5×5-fold cross-validation (5×5-CV) is used to generate training datasets and testing datasets. The corresponding experimental results are evaluated by using the following F1 score and G-mean [35] in the form of "average ± standard deviation":

$$F_1 = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{28}$$

$$G-mean = \sqrt{ACC_{min} \times ACC_{maj}} \tag{29}$$

where F_1 denotes the hybrid indicator of the recall and precision of a classifier; *G-mean* denotes the geometric means of the accuracy of a classifier for minority and majority classes. The corresponding formulas of recall, precision, and the accuracy of minority and majority classes are as follows

$$ACC_{min} = Recall = \frac{TP}{TP + FN} \tag{30}$$

$$Precison = \frac{TP}{TP + FP} \tag{31}$$

$$ACC_{maj} = \frac{TN}{TN + FP} \tag{32}$$

where *TP*, *FN*, *FP*, *TN* are the four indicators obtained from a confusion matrix, as shown in Table 2.

Table 3. Confusion matrix under imbalanced classification problems

Classes	Predictive minority class	Predictive majority class
Actual minority class	*TP*	*FN*
Actual majority class	*FP*	*TN*

Moreover, the following options are considered to construct different variants of the EBRB decision model for the purpose of performance comparison:

(1) For the process of constructing EBRB, there are four kinds of options: none of data oversampling method is used to balance imbalanced datasets, denoted as C_{none}, and the use of ADASYN [6], Borderline-SMOTE[5], and Safe-Level-SMOTE [4] techniques, denoted as C_{ADAS}, C_{BORD}, and C_{SAFE}, to balance imbalanced datasets, respectively; Noting that all the three data oversampling techniques are used to generate diverse base EBRBs for the ensemble EBRB decision model.

(2) For the process of ensemble inference, an average-based integrated belief degree is defined to produce the final output of the ensemble EBRB decision model, denoted as I_{avg}, where the average-based integrated belief degree is written as follows:

$$\beta_n = \frac{\sum_{s=1}^{S} \beta_n^s}{S} \tag{33}$$

Additionally, for the construction of ensemble and traditional EBRB decision models, the parameter configuation is defined as follows: 1) the number of reference values for each antecedent attribute is set as 5 and the number of consequents is equal to the number of classes; 2) the activation factor λ in Eq. (12) is set as 1; 3) the attribute weights used in traditional EBRB decision model are all defined as 1; and 4) the ensemble EBRB decision model is defined to have 30 base EBRBs, which means that C_{ADAS}, C_{BORD}, and C_{SAFE} are evenly used to generate 10 base EBRBs.

5.2. Comparison of ensemble and traditional EBRB decision model

In this subsection, the proposed ensemble EBRB decision model is evaluated against the EBRB decision models equipped with the four options detailed in Section 5.1. Tables 4 and 5 show the experimental results of 26 imbalanced datasets in the terms of F1 score and G-mean.

Table 4. Comparison of average F1 score for traditional and ensemble EBRB decision models

Dataset	Traditional EBRB decision model				Ensemble EBRB decision model
	C_{none}	C_{ADAS}	C_{BORD}	C_{SAFE}	
Eco-1	**84.17±0.00(1)**	76.21±0.71(5)	76.52±0.29(4)	79.13±0.52(2)	77.23±0.29(3)
Eco-2	86.84±0.00(2)	63.32±0.16(5)	74.67±0.38(4)	**87.50±0.80(1)**	77.21±0.35(3)
Eco-3	60.94±0.00(4)	55.68±0.33(5)	63.39±0.24(3)	**70.04±1.28(1)**	64.06±0.27(2)
Eco-4	79.76±0.00(3)	45.75±0.94(5)	64.05±1.03(4)	**83.91±2.31(1)**	82.11±0.00(2)
Eco-5	79.81±0.00(2.5)	65.05±1.09(5)	74.33±1.33(4)	79.81±0.00(2.5)	**83.20±0.85(1)**
Eco-6	74.51±0.00(3)	55.34±1.40(3)	71.18±2.10(4)	75.35±1.33(2)	**80.94±0.89(1)**
Eco-7	75.78±0.00(2)	64.40±1.38(5)	66.91±0.00(4)	74.78±0.82(3)	**81.83±0.71(1)**
Eco-8	**79.84±0.00(1)**	54.72±1.73(5)	66.35±0.06(4)	77.71±1.23(2)	72.85±1.45(3)
Eco-9	78.57±0.00(2)	62.12±2.27(5)	69.19±1.05(4)	77.24±1.09(3)	**80.63±1.46(1)**
Eco-10	71.00±0.00(3)	54.75±1.16(5)	64.47±1.00(4)	74.68±0.58(2)	**76.43±1.76(1)**
Eco-11	**78.36±0.00(1)**	51.39±0.44(5)	66.65±0.63(4)	76.63±0.44(3)	76.81±1.18(2)
Yea-1	33.09±0.00(5)	56.51±0.22(4)	57.01±0.25(3)	57.11±0.44(2)	**57.30±0.20(1)**
Yea-2	18.29±0.00(5)	49.95±0.14(4)	51.91±0.00(3)	65.77±0.58(2)	**69.94±0.32(1)**
Yea-3	16.41±0.00(5)	26.67±0.12(4)	33.25±0.41(2)	29.21±1.01(3)	**39.91±0.81(1)**
Yea-4	37.48±0.00(2)	21.29±0.15(5)	29.94±0.07(4)	**52.24±1.00(1)**	36.71±0.39(3)
Yea-5	22.86±0.00(4)	21.12±0.28(5)	27.29±0.59(3)	29.84±3.05(2)	**45.72±2.13(1)**
Yea-6	64.60±0.00(4)	61.54±0.87(5)	71.67±0.48(3)	74.87±0.73(2)	77.25±0.81(1)
Yea-7	63.81±0.00(2.5)	16.82±0.26(5)	37.60±0.65(4)	63.81±0.00(2.5)	**65.15±0.67(1)**
Yea-8	**82.82±0.00(1)**	44.89±0.21(5)	64.49±0.37(4)	80.85±0.29(2)	78.16±0.62(3)
Yea-9	55.63±0.00(3)	37.03±0.75(5)	49.95±0.35(4)	60.75±1.03(2)	**60.88±0.60(1)**
Yea-10	33.13±0.00(4)	33.39±0.86(3)	30.53±0.41(5)	35.51±1.69(2)	**40.91±1.76(1)**
Veh-1	40.51±0.00(5)	**53.44±1.35(1)**	52.89±0.65(3)	49.43±1.51(4)	52.90±0.73(2)
Veh-2	42.89±0.00(5)	53.94±1.09(2)	54.47±0.31(3)	52.12±0.82(4)	**55.29±0.61(1)**
New-1	76.64±0.00(5)	81.25±0.67(4)	88.54±0.90(2)	86.78±0.85(3)	**91.63±1.71(1)**
New-2	78.34±0.00(5)	81.12±0.64(4)	86.55±0.69(3)	88.04±0.85(2)	**90.31±1.04(1)**
Pim	53.85±0.00(5)	62.54±0.23(3)	60.15±0.67(4)	**63.41±0.41(1)**	63.21±0.37(2)
Average rank	3.27	4.31	3.58	2.19	**1.58**

Table 5. Comparison of average G-mean for traditional and ensemble EBRB decision models

Dataset	Traditional EBRB decision model				Ensemble EBRB decision model
	C_{none}	C_{ADAS}	C_{BORD}	C_{SAFE}	
Eco-1	**88.63±0.00(2)**	88.41±0.48(4)	87.79±0.10(5)	88.77±0.39(1)	88.49±0.17(3)
Eco-2	92.79±0.00(2)	86.52±0.10(5)	91.06±0.12(4)	**94.42±0.19(1)**	91.85±0.08(3)
Eco-3	71.70±0.00(5)	87.39±0.11(4)	89.72±0.07(2)	88.45±0.78(3)	**89.89±0.07(1)**
Eco-4	83.29±0.00(5)	88.22±0.26(4)	91.76±0.13(3)	91.94±1.44(2)	**93.75±0.00(1)**
Eco-5	85.22±0.00(4.5)	88.08±1.03(2)	86.87±1.24(3)	85.22±0.00(4.5)	**90.83±0.09(1)**
Eco-6	82.82±0.00(4)	82.49±0.89(5)	86.23±1.98(2)	84.34±0.88(3)	**90.70±0.19(1)**
Eco-7	84.00±0.00(3)	85.51±0.31(2)	82.73±0.00(5)	83.88±0.10(4)	**90.93±0.10(1)**
Eco-8	88.42±0.00(2)	86.26±1.11(5)	**90.54±0.00(1)**	88.23±0.11(3)	87.70±0.15(4)
Eco-9	85.01±0.00(3)	85.51±1.96(2)	83.72±0.12(5)	84.84±0.14(4)	**90.13±0.24(1)**
Eco-10	78.35±0.00(5)	81.23±0.99(3)	80.86±0.19(4)	82.62±0.09(2)	**85.58±1.51(1)**
Eco-11	87.47±0.00(2)	83.51±0.17(5)	86.88±0.20(4)	87.22±0.06(3)	**87.95±1.04(1)**
Yea-1	45.62±0.00(5)	66.74±0.22(4)	**68.94±0.24(1)**	69.54±0.44(3)	69.71±0.16(2)
Yea-2	31.07±0.00(5)	86.08±0.07(3)	86.24±0.02(2)	73.37±0.38(4)	**92.50±0.14(1)**
Yea-3	27.59±0.00(5)	**84.75±0.06(1)**	84.37±0.43(2)	42.85±0.96(4)	81.06±0.46(3)
Yea-4	51.87±0.00(5)	86.13±0.07(3)	**89.21±0.02(1)**	69.30±0.87(4)	86.17±0.75(2)
Yea-5	32.66±0.00(5)	**67.23±0.31(1)**	64.89±0.84(2)	40.42±2.39(4)	62.25±1.77(3)
Yea-6	72.94±0.00(5)	**89.41±0.53(1)**	89.03±0.08(2)	82.77±0.57(4)	85.27±0.09(3)
Yea-7	72.68±0.00(2.5)	71.64±0.30(4)	70.80±0.07(5)	72.68±0.00(2.5)	**72.77±0.05(1)**
Yea-8	88.15±0.00(4)	83.12±0.14(5)	89.56±0.09(2)	88.40±0.04(3)	**89.95±0.51(1)**
Yea-9	65.93±0.00(5)	72.83±0.84(4)	76.81±0.28(2)	73.39±0.87(3)	**78.02±0.43(1)**
Yea-10	46.46±0.00(5)	**72.20±0.92(1)**	66.60±0.56(3)	50.29±1.39(4)	67.55±0.71(2)
Veh-1	56.05±0.00(5)	**69.03±1.21(1)**	68.24±0.56(4)	66.51±1.27(3)	68.32±0.64(2)
Veh-2	57.92±0.00(5)	70.39±0.51(2)	69.90±0.21(3)	67.00±0.80(4)	**70.57±0.50(1)**
New-1	78.99±0.00(5)	95.26±0.67(2)	**97.15±0.26(1)**	87.80±0.72(4)	95.11±1.26(3)
New-2	80.53±0.00(5)	95.07±0.22(3)	**96.50±0.22(1)**	88.87±0.72(4)	96.06±1.12(2)
Pim	63.29±0.00(5)	70.66±0.13(3)	68.80±0.57(4)	**71.37±0.36(1)**	71.27±0.31(2)
Average rank	4.19	3.04	2.81	3.15	**1.81**

From Tables 4 and 5, it can be seen that there are some substantial improvements for the proposed ensemble EBRB decision model over the traditional EBRB decision models, which is in line with the conception that ensemble learning is applicable for improving the performance of EBRB decision model under imbalanced classification problems. Specifically, two preliminary conclusions are drawn based on F1 score and G-mean:

(1) In the comparison of F1 score, the proposed ensemble EBRB decision model obtains the best F1 score in 15 of the 26 datasets and its average rank is 1.75. The traditional EBRB decision model with C_{SAFE} obtains the second best F1 score, owing to its cautious sampling process that only synthesizes the data with minority class in the safety region far away from the borderline. As for the traditional EBRB decision model with C_{BORD} and C_{ADAS}, they all focus on the borderline area to weaken the dominance of the data with majority class in overlapping area. Thus, the error of the data with minority class can be increased, leading to a rather inferior performance on F1 score. As for the ensemble EBRB decision model which is constructed by three data oversampling methods, its performance can be improved significantly better than the traditional EBRB decision models with different options.

(2) In the comparison of G-mean, Table 5 shows that any one of the three data oversampling methods can bring

considerable improvements compared to the traditional EBRB decision model without data oversampling. For the traditional EBRB decision model with C_{ADAS}, it obtains the second-best G-mean. This means that this data oversampling method has advantage in increasing the recall rate of EBRB decision model with a relatively low cost of decreasing the specificity. As for the traditional EBRB decision model with C_{SAFE}, it is useful to sample the data with minority class in the safe region for guaranteeing preferable precision, but the recall rate may not get sufficient improvements, namely, a small improvement in G-mean. Owing to the diverse data oversampling methods and consistency-based combination methods, the proposed ensemble EBRB decision model has the highest G-mean in almost half of the selected datasets.

In summary, the proposed ensemble EBRB decision model can handle imbalanced classification problems better than other traditional EBRB decision models in some selected imbalanced datasets. To further evaluate the effectiveness of the proposed ensemble EBRB decision model the next section aims to compare ensemble EBRB decision models in the respect of different component generation and combination methods.

5.3. Comparison of different construction and inference processes for ensemble EBRB decision model

In this subsection, the construction and inference options detailed in Section 5.1 are investigated to illustrate the effectiveness of the proposed ensemble EBRB decision model. Tables 6 to 7 show the experimental results regarding F1 score and G-mean. Noting that each kind of ensemble EBRB decision model shown in Tables 6 to 7 has only one difference comparing to the proposed ensemble EBRB decision model, including the ensemble EBRB decision model with C_{none} means that the construction of base EBRBs without data oversampling techniques, the ensemble EBRB decision model with C_{ADAS}, C_{BORD}, or C_{SAFE} means that the construction of base EBRBs using only one kind of data oversampling technique, and the ensemble EBRB decision model with I_{avg} means that the average-based integrated belief degree is used to replace the consistency-based integrated belief degree in the inference process of ensemble EBRB decision model.

Table 6. Comparison of average F1 score for different ensemble EBRB decision models

Dataset	Ensemble EBRB decision model					
	C_{none}	C_{ADAS}	C_{BORD}	C_{SAFE}	I_{avg}	This paper
Eco-1	**81.52±0.77(1)**	76.19±0.75(6)	76.25±0.39(5)	77.98±0.20(2)	76.69±0.30(4)	77.23±0.29(3)
Eco-2	**87.51±0.00(1)**	70.66±0.63(6)	74.84±0.66(5)	85.43±0.46(2)	77.83±0.65(3)	77.21±0.35(4)
Eco-3	65.68±1.62(2)	55.95±0.28(6)	63.37±0.78(5)	**70.17±1.15(1)**	63.91±0.33(4)	64.06±0.27(3)
Eco-4	**84.14±1.87(1)**	52.71±2.03(6)	75.59±0.94(5)	82.06±0.00(3)	79.71±1.20(4)	82.11±0.00(2)
Eco-5	77.60±2.10(5)	72.15±0.96(6)	80.06±1.38(4)	81.65±1.32(3)	82.89±0.97(2)	**83.20±0.85(1)**
Eco-6	74.29±2.69(5)	67.81±1.98(6)	77.33±1.04(4)	78.78±5.87(3)	80.82±1.47(2)	**80.94±0.89(1)**
Eco-7	79.70±0.89(4)	74.16±1.61(6)	75.80±2.27(5)	**84.62±3.89(1)**	81.16±0.84(3)	81.83±0.71(2)
Eco-8	79.41±1.35(2)	64.27±1.30(6)	78.13±1.59(3)	**80.87±1.24(1)**	72.27±1.09(5)	72.85±1.45(4)
Eco-9	79.59±1.74(4)	65.89±1.81(6)	79.26±2.08(5)	80.30±2.80(2)	80.17±1.68(3)	**80.63±1.46(1)**
Eco-10	78.23±1.96(2)	64.81±2.57(6)	70.88±1.51(5)	**79.30±1.07(1)**	75.08±1.49(4)	76.43±1.76(3)
Eco-11	**79.82±0.73(1)**	64.00±3.29(6)	75.16±1.03(5)	76.38±1.32(4)	77.38±1.18(2)	76.81±1.18(3)
Yea-1	31.26±0.14(6)	56.00±0.10(5)	**57.51±0.35(1)**	56.60±0.42(4)	57.40±0.38(2)	57.30±0.20(3)
Yea-2	5.88±0.00(6)	50.74±0.13(4)	54.16±0.28(3)	48.73±0.68(5)	67.34±0.32(2)	**69.94±0.32(1)**
Yea-3	11.82±1.74(6)	32.06±0.27(4)	33.89±0.43(3)	12.56±1.59(5)	37.23±0.38(2)	**39.91±0.81(1)**
Yea-4	25.34±0.27(6)	25.76±0.34(5)	32.27±0.17(4)	**46.41±2.73(1)**	34.92±0.32(3)	36.71±0.39(2)
Yea-5	19.89±2.85(6)	24.22±0.43(5)	34.79±3.32(3)	24.95±2.85(4)	40.37±2.91(2)	**45.72±2.13(1)**

Yea-6	65.74±1.30(6)	68.00±1.16(5)	71.95±1.68(4)	73.19±0.37(3)	76.43±0.78(2)	**77.25±0.81(1)**
Yea-7	64.72±1.52(5)	23.26±0.77(6)	64.81±0.82(4)	**65.75±0.53(1)**	65.07±0.65(3)	65.15±0.67(2)
Yea-8	**81.31±0.30(1)**	61.87±0.18(6)	69.93±0.16(5)	81.23±0.24(2)	77.45±0.77(4)	78.16±0.62(3)
Yea-9	53.43±1.61(5)	51.91±0.54(6)	59.06±0.58(4)	60.23±0.40(3)	**61.57±0.47(1)**	60.88±0.60(2)
Yea-10	31.26±0.94(6)	33.53±0.62(4)	34.83±0.48(3)	33.09±0.00(5)	38.98±1.18(2)	**40.91±1.76(1)**
Veh-1	41.77±1.13(6)	**55.03±0.13(1)**	50.94±0.66(4)	49.47±0.26(5)	52.15±0.83(3)	52.90±0.73(2)
Veh-2	40.63±2.24(6)	**56.89±0.42(1)**	52.55±0.98(4)	52.38±0.81(5)	55.31±0.63(2)	55.29±0.61(3)
New-1	72.73±0.00(6)	75.39±0.79(5)	86.44±1.03(3)	78.67±0.85(4)	90.54±0.60(2)	**91.63±1.71(1)**
New-2	71.55±0.00(6)	77.78±1.03(4)	85.26±1.37(3)	76.58±0.72(5)	90.18±0.75(2)	**90.31±1.04(1)**
Pim	52.16±0.54(6)	**64.28±0.44(1)**	61.91±0.31(5)	63.46±0.46(2)	62.67±0.35(4)	63.21±0.37(3)
Average rank	4.27	4.92	4.00	2.96	2.77	**2.08**

From Table 6, the ensemble EBRB decision model proposed in this paper achieved first place with an average rank of 2.21, ahead of the runner-up (I_{avg}) 0.65. C_{SAFE} was third, but the average rank of it is close to I_{avg}. Based on the performances in the F1 score of the different kinds on ensemble EBRB decision models on each dataset and the average ranks, several interesting issues can be concluded as follows:

(1) It is not difficult to find out that C_{ADAS}, C_{BORD} and C_{SAFE} have a distinguished performance in F1 score on the listed datasets according to their average ranks. C_{SAFE} has superior performance over its peers. The reason behind this phenomenon is probably that C_{SAFE} only oversamples the minority examples far from the class boundary, leading to a rather good performance of ensemble EBRB decision model generated by C_{SAFE} in precision and thus a higher value in F1. However, none of the above three ensemble schemes outperform the proposed method, in other words, owing to the combination of C_{ADAS}, C_{BORD}, and C_{SAFE} simultaneously, the proposed EBRB decision model is able to provide an advanced performance because of the diversity enhancement of the base EBRB decision models.

(2) Noting that the proposed EBRB decision model outperformed other EBRB decision models on 10 listed datasets while I_{avg}, which utilized average-based instead of consistency-based integration belief degree in the inference process, only obtains the first place once on Yea-9. The results from the cases investigated indicates that the proposed consistency-based ensemble inference method brings considerable improvements for ensemble EBRB decision model to handle class imbalance. By distinguishing the reliabilities of each component's outputs, it can be found that inconsistent activated rules have few contributions to the final outputs. From Table 6, although C_{SAFE} gets the best result in six datasets, its average rank is still lower than that of I_{avg}, which implies that the diversity-based base EBRB decision model generation method plays a more fundamental role in learning from the imbalanced datasets while diversity-based ensemble inference strategy can be treated as an enhancement method to improve the performance of ensemble EBRB decision model.

Table 7. Comparison of average G-mean for different ensemble EBRB decision models

Dataset	Ensemble EBRB decision model					
	C_{none}	C_{ADAS}	C_{BORD}	C_{SAFE}	I_{avg}	This paper
Eco-1	87.07±0.58(6)	88.32±0.45(3)	88.00±0.20(5)	**88.84±0.11(1)**	88.20±0.16(4)	88.49±0.17(2)
Eco-2	92.10±0.00(2)	89.76±0.21(6)	91.11±0.21(5)	**93.40±0.35(1)**	92.03±0.18(3)	91.85±0.08(4)
Eco-3	75.97±1.19(6)	87.35±0.06(5)	89.74±0.19(4)	89.84±0.47(3)	89.86±0.09(2)	89.89±0.07(1)
Eco-4	87.24±1.55(6)	91.63±0.54(4)	**95.46±0.15(1)**	88.85±0.00(5)	93.53±0.11(3)	93.75±0.00(2)
Eco-5	81.39±1.70(6)	89.58±0.12(4)	90.50±0.15(3)	85.49±1.27(5)	90.78±0.09(2)	**90.83±0.09(1)**
Eco-6	79.41±2.03(6)	88.04±0.33(4)	90.25±0.16(3)	83.75±3.98(5)	90.31±0.86(2)	**90.70±0.19(1)**
Eco-7	84.39±0.10(6)	89.68±0.30(3)	89.10±1.58(4)	88.89±2.74(5)	90.30±1.18(2)	**90.93±0.10(1)**
Eco-8	85.76±1.11(6)	91.85±0.23(2)	**93.27±1.03(1)**	88.53±0.11(3)	87.64±0.11(5)	87.70±0.15(4)

Eco-9	86.74±1.42(6)	87.60±0.39(4)	88.98±1.52(3)	87.37±1.34(5)	90.09±0.26(2)	**90.13±0.24(1)**
Eco-10	83.44±1.55(5)	84.86±0.63(3)	82.54±1.44(6)	84.27±1.01(4)	84.97±1.22(2)	**85.58±1.51(1)**
Eco-11	87.64±0.08(4)	87.65±1.17(3)	87.41±0.78(5)	87.38±0.17(6)	88.74±0.86(2)	**87.95±1.04(1)**
Yea-1	44.02±1.22(6)	66.44±0.07(5)	**69.87±0.25(1)**	68.48±0.37(4)	69.83±0.33(2)	69.71±0.16(3)
Yea-2	15.49±0.00(6)	86.90±0.08(4)	88.13±0.16(3)	57.82±0.60(5)	92.30±0.21(2)	**92.50±0.14(1)**
Yea-3	21.20±3.09(6)	83.49±0.45(2)	**83.57±0.42(1)**	26.58±2.95(5)	81.87±0.38(3)	81.06±0.46(4)
Yea-4	36.44±0.00(6)	88.08±0.11(2)	**89.19±0.67(1)**	60.98±2.99(5)	87.31±0.05(3)	86.17±0.75(4)
Yea-5	29.44±4.00(6)	**70.61±0.62(1)**	62.14±2.70(3)	35.27±2.51(5)	61.18±2.41(4)	62.25±1.77(2)
Yea-6	73.58±0.66(6)	**88.65±0.46(1)**	86.33±0.25(2)	80.66±0.44(5)	85.41±0.60(3)	85.27±0.09(4)
Yea-7	72.14±1.27(6)	**77.02±2.37(1)**	72.74±0.05(4)	72.69±0.02(5)	72.75±0.04(3)	72.77±0.05(2)
Yea-8	86.79±0.28(6)	88.82±0.15(4)	**90.67±0.28(1)**	87.67±0.25(5)	90.13±0.73(2)	89.95±0.51(3)
Yea-9	63.75±1.29(6)	**78.32±0.33(1)**	77.74±0.38(4)	73.73±0.28(5)	78.15±0.52(2)	78.02±0.43(3)
Yea-10	44.31±0.80(6)	**72.93±0.64(1)**	70.87±0.66(2)	45.90±0.00(5)	68.97±0.95(3)	67.55±0.71(4)
Veh-1	56.18±0.95(6)	**70.40±0.11(1)**	66.47±0.58(4)	65.09±0.24(5)	67.69±0.71(3)	68.32±0.64(2)
Veh-2	54.17±1.89(6)	**72.30±0.35(1)**	67.76±0.58(4)	67.43±0.72(5)	70.50±0.53(3)	70.57±0.50(2)
New-1	75.59±0.00(6)	93.32±0.29(4)	**96.77±0.29(1)**	80.59±0.72(5)	95.51±0.13(2)	95.11±1.26(3)
New-2	74.94±0.00(5)	93.87±0.66(3)	96.03±0.75(2)	79.07±0.64(4)	94.93±0.73(3)	**96.06±1.12(1)**
Pim	61.38±0.42(6)	**72.08±0.35(1)**	70.22±0.25(5)	71.44±0.38(2)	70.82±0.29(4)	71.27±0.31(3)
Average rank	5.69	2.81	3.00	4.35	2.73	**2.31**

From Table 7, it can be seen that the proposed ensemble EBRB decision model and its variant I_{avg} outperform their rivals again while C_{SAFE} only get the second worst place which is just better than C_{none} in G-mean. Two preliminary conclusions can be drawn:

(1) When G-mean is considered as a performance measure, C_{ADAS} outperforms C_{SAFE} and C_{BORD}, this is mainly because C_{ADAS} focuses on the borderline area and aims to rebalance each hard-to-learn area detected by the *k*NN algorithm to degrade the dominance of majority class. However, as the experimental results shown in Table 5, when C_{ADAS} is applied on the single EBRB decision model, the performance is still inferior. The reason is that C_{ADAS} may increase the risk of overfitting since the synthetic positive sample generation and hard-to-learn area detection process, leading to an inferior performance in G-mean. When C_{ADAS} and EBRB decision model are used to develop an ensemble model, the generation ability is improved thus enabling it to provide more robust prediction results.

(2) Since the proposed ensemble EBRB decision model gets the best result in both G-mean and F1 score comparing to its variants, one thing can be guaranteed is that the diversity-based base EBRB construction method and diversity-based ensemble inference method are effective to help EBRB decision model learn from the imbalanced datasets, in which the former method let the ensemble EBRB decision model have more diverse components comparing to the EBRB decision models which combines with one kind of oversampling technique; the latter method can enhance the robustness of prediction results as they can be treated as the perturbation in the combination process.

In summary, the effectiveness of diversity-based base EBRB construction method is demonstrated by the comparison result that the ensemble EBRB decision model outperforms the ensemble EBRB decision model only considering one kind of oversampling technique. For the proposed consistency-based ensemble inference, it can be treated as an enhancement method to strengthen the robustness of the unified prediction by distinguishing the reliabilities of each component output.

5.4. Comparison of ensemble EBRB decision models with and without attribute weights calculation

In this subsection, the proposed attribute weight calculation method is investigated to demonstrate the effectiveness of the proposed ensemble EBRB decision model. For this propose, the ensemble EBRB decision model with the average- based integrated belief degree is used as a baseline classifier. Tables 8 and 9 show the experimental results regarding F1 score and G-mean.

Table 8. Comparison of F1 score for ensemble EBRB decision models with or without attribute weights calculation

Dataset	Without attribute weight calculation		With attribute weight calculation	
	Ensemble EBRB+I_{avg}	Ensemble EBRB	Ensemble EBRB+I_{avg}	Ensemble EBRB
Eco-1	**77.31±0.50(1.5)**	**77.31±0.50(1.5)**	76.69±0.30(4)	77.23±0.29(3)
Eco-2	77.45±0.54(3)	**81.07±0.44(1)**	77.83±0.65(2)	77.21±0.35(4)
Eco-3	64.71±0.70(2)	**64.81±0.49(1)**	63.91±0.33(4)	64.06±0.27(3)
Eco-4	79.42±1.46(4)	79.53±0.71(3)	79.71±1.20(2)	**82.11±0.00(1)**
Eco-5	81.16±1.86(4)	81.49±1.63(3)	82.89±0.97(2)	**83.20±0.85(1)**
Eco-6	78.18±2.35(4)	80.50±3.16(3)	80.82±1.47(2)	**80.94±0.89(1)**
Eco-7	77.30±2.37(4)	80.11±1.41(3)	81.16±0.84(2)	**81.83±0.71(1)**
Eco-8	72.32±0.62(3)	72.34±0.41(2)	72.27±1.09(4)	**72.85±1.45(1)**
Eco-9	74.25±2.32(4)	74.30±2.58(3)	80.17±1.68(2)	**80.63±1.46(1)**
Eco-10	71.81±0.93(3)	70.01±1.71(4)	75.08±1.49(2)	**76.43±1.76(1)**
Eco-11	74.66±0.54(4)	75.11±0.44(3)	**77.38±1.18(1)**	76.81±1.18(2)
Yea-1	57.08±0.29(3)	57.05±0.47(4)	**57.40±0.38(1)**	57.30±0.20(2)
Yea-2	67.18±0.48(4)	67.53±0.34(2)	67.34±0.32(3)	**69.94±0.32(1)**
Yea-3	37.12±0.60(4)	37.33±0.33(2)	37.23±0.38(3)	**39.91±0.81(1)**
Yea-4	36.56±0.58(3)	**37.23±0.33(1)**	34.92±0.32(4)	36.71±0.39(2)
Yea-5	37.81±2.66(4)	43.40±2.53(2)	40.37±2.91(3)	**45.72±2.13(1)**
Yea-6	76.58±0.45(2)	76.39±0.34(4)	76.43±0.78(3)	**77.25±0.81(1)**
Yea-7	63.13±1.33(4)	63.81±0.00(3)	65.07±0.65(2)	**65.15±0.67(1)**
Yea-8	76.29±0.38(3)	77.24±0.27(3)	77.45±0.77(2)	**78.16±0.62(1)**
Yea-9	60.60±0.49(4)	60.80±0.23(3)	**61.57±0.47(1)**	60.88±0.60(2)
Yea-10	40.62±0.94(3)	**43.81±1.85(1)**	38.98±1.18(4)	40.91±1.76(2)
Veh-1	50.99±0.73(4)	51.21±0.91(3)	52.15±0.83(2)	**52.90±0.73(1)**
Veh-2	52.82±0.70(4)	53.06±0.73(3)	**55.31±0.63(1)**	55.29±0.61(2)
New-1	95.14±0.00(2)	**96.17±0.00(1)**	90.54±0.60(4)	91.63±1.71(3)
New-2	94.00±0.98(2)	**94.93±0.55(1)**	90.18±0.75(4)	90.31±1.04(3)
Pima	62.45±0.33(3)	62.39±0.60(4)	62.67±0.35(2)	**63.21±0.37(1)**
Average rank	3.29	2.48	2.54	**1.65**

Table 9. Comparison of G-mean for ensemble EBRB decision models with or without attribute weights calculation

Dataset	Without attribute weight calculation		With attribute weight calculation	
	Ensemble EBRB+I_{avg}	Ensemble EBRB	Ensemble EBRB+I_{avg}	Ensemble EBRB
Eco-1	88.21±0.39(2.5)	88.21±0.39(2.5)	88.20±0.16(4)	**88.49±0.17(1)**
Eco-2	91.93±0.13(3)	**92.89±0.13(1)**	92.03±0.18(2)	91.85±0.08(4)
Eco-3	89.96±0.17(2)	**90.05±0.13(1)**	89.86±0.09(4)	89.89±0.07(3)
Eco-4	93.49±0.14(3)	93.41±0.08(4)	93.53±0.11(2)	**93.75±0.00(1)**
Eco-5	89.65±1.30(4)	90.19±1.06(3)	90.78±0.09(2)	**90.83±0.09(1)**
Eco-6	88.89±0.97(4)	89.25±1.09(3)	90.31±0.86(2)	**90.70±0.19(1)**
Eco-7	88.19±1.37(4)	89.63±1.47(3)	90.30±1.18(2)	**90.93±0.10(1)**
Eco-8	87.64±0.06(2.5)	87.54±0.06(4)	87.64±0.11(2.5)	**87.70±0.15(1)**
Eco-9	85.43±1.59(4)	86.89±1.82(3)	90.09±0.26(2)	**90.13±0.24(1)**

Eco-10	83.60±0.73(3)	82.26±0.87(4)	84.97±1.22(2)	**85.58±1.51(1)**
Eco-11	86.96±0.07(4)	87.00±0.06(3)	**88.74±0.86(1)**	87.95±1.04(2)
Yea-1	69.49±0.23(4)	69.48±0.39(3)	**69.83±0.33(1)**	69.71±0.16(2)
Yea-2	92.17±0.28(3)	89.63±0.13(4)	92.30±0.21(3)	**92.50±0.14(1)**
Yea-3	81.86±0.60(2)	79.48±0.88(4)	**81.87±0.38(1)**	81.06±0.46(3)
Yea-4	**87.54±0.09(1)**	85.85±0.03(4)	87.31±0.05(2)	86.17±0.75(3)
Yea-5	62.32±1.08(2)	**62.60±1.88(1)**	61.18±2.41(4)	62.25±1.77(3)
Yea-6	**86.04±0.06(1)**	85.46±0.48(2)	85.41±0.60(3)	85.27±0.09(4)
Yea-7	72.65±0.06(4)	72.68±0.00(3)	72.75±0.04(2)	**72.77±0.05(1)**
Yea-8	88.96±0.25(4)	89.09±0.24(3)	**90.13±0.73(1)**	89.95±0.51(2)
Yea-9	77.76±0.36(4)	78.14±0.46(2)	**78.15±0.52(1)**	78.02±0.43(3)
Yea-10	68.68±0.53(3)	**69.19±1.38(1)**	68.97±0.95(2)	67.55±0.71(4)
Veh-1	66.45±0.67(4)	66.62±0.79(3)	67.69±0.71(2)	**68.32±0.64(1)**
Veh-2	68.21±0.63(4)	68.43±0.61(3)	70.50±0.53(2)	**70.57±0.50(1)**
New-1	98.89±0.00(2)	**99.16±0.00(1)**	95.51±0.13(3)	95.11±1.26(4)
New-2	98.48±0.56(2)	**98.70±0.51(1)**	94.93±0.73(4)	96.06±1.12(3)
Pima	70.60±0.28(3)	70.56±0.49(4)	70.82±0.29(2)	**71.27±0.31(1)**
Average rank	3.04	2.71	2.25	**2.04**

As Tables 8 and 9 show, the proposed attribute weight calculation method is helpful to improve the performance of the ensemble EBRB decision models in both F1 score and G-mean. Preliminary conclusions are drawn based on the experimental results:

(1) For F1 score, the ensemble EBRB decision model with attribute weights calculation get the best result in 19 datasets, which is about 68 percentages of all the given datasets. Apart from New-1, New-2, Veh-1, and Veh-3, the ensemble EBRB decision model with attribute weight calculation can achieve better performance in F1 score, especially for Eco-5, Eco-9, Eco-10; For G-mean, the ensemble EBRB decision model with attribute weight calculation has the best result on 18 datasets. Additionally, the attribute weight calculation method is more effective when the given datasets have median scale of features but it is low effective for large or small scale of features because of dimensionality curse and over-fitting problems.

(2) From Table 9, the ensemble EBRB + I_{avg} with attribute weight calculation has a better average rank than that without attribute weight calculation, which can manifest the effectiveness of attribute weights calculation. Apart from the diversity enhancement of each base EBRB, this is because the information gain-based attribute weight calculation is useful to highlight the relevant attributes in rule activation process. Besides, comparing to the feature selection methods proposed for the conventional EBRB decision model, the reason why the attribute weight calculation works properly is because the calculation process of attribute weights comes after the class rebalancing process thus alleviating or even eliminating the bias of conventional feature engineering techniques to the majority class.

In summary, the comparative study demonstrates that the attribute weight calculation method based on the information gain can bring significant improvements for ensemble EBRB decision model by increasing the diversity of components and alleviating the impacts of class overlap.

5.5. Comparative analysis with ensemble EBRB decision model and conventional classifiers

To further verify the validity of the proposed ensemble EBRB decision model, some commonly used machine learning-based classifiers, including support vector machine (SVM), decision tree (DT), artificial neural network (ANN), and *k*- nearest neighbor (KNN), are used as the baseline component to construct ensemble classifiers, where SVM is defined to use radial basis function (RBF) as the kernel function and the reciprocal of the attribute in dataset as coefficient, and the penalty coefficient is set as 0.5; DT is defined to use Gini coefficient to choose split feature and the maximum number of leaf nodes is not limited; ANN is defined to use ReLU as activation function and each hidden layer has 10 neurons, the maximum number of iterations is set to 1000; KNN is defined to use Euclidean distance to measure the similarity between data under k=5. For the sake of fairness, Step 1 and Step 2 in Section 4.2 will also be applied to the above baselines to construct the ensemble models. The ensemble classifiers are abbreviated as E-SVM, E-DT, E-ANN, and E-KNN, respectively. Tables 10 and 11 show the experimental results of F1 scores and G-mean.

Table 10. Comparison of average F1 score for ensemble EBRB decision models with ensemble classifiers

Dataset	EBRB	E-SVM	E-DT	E-MLP	E-KNN	Ensemble EBRB
Eco-1	**84.17±0.00(1)**	77.96±0.49(4)	78.44±0.69(3)	76.39±0.64(6)	79.50±0.66(2)	77.23±0.29(5)
Eco-2	86.84±0.00(2)	**88.14±0.42(1)**	82.28±0.45(4)	81.92±1.33(5)	84.44±0.47(3)	77.21±0.35(6)
Eco-3	60.94±0.00(6)	65.29±1.90(2)	64.19±1.06(3)	**66.94±1.17(1)**	62.53±1.27(5)	64.06±0.27(4)
Eco-4	79.76±0.00(3)	81.72±2.84(2)	76.91±4.05(4)	76.76±1.59(5)	76.57±2.31(6)	**82.11±0.00(1)**
Eco-5	79.81±0.00(5)	**87.11±0.94(1)**	74.28±1.30(6)	81.34±1.89(4)	82.10±1.79(3)	83.20±0.85(2)
Eco-6	74.51±0.00(5)	75.97±1.31(4)	74.26±4.41(6)	**82.23±1.07(1)**	77.13±0.69(3)	80.94±0.89(2)
Eco-7	75.78±0.00(6)	84.02±0.95(2)	80.04±1.49(5)	80.70±2.61(4)	**86.30±1.65(1)**	81.83±0.71(3)
Eco-8	79.84±0.00(3)	79.96±2.39(2)	**85.56±1.68(1)**	79.81±1.68(4)	77.97±1.40(5)	72.85±1.45(6)
Eco-9	78.57±0.00(6)	**83.73±1.68(1)**	82.14±2.52(2)	81.43±2.52(3)	80.34±2.28(5)	80.63±1.46(4)
Eco-10	71.00±0.00(3)	71.21±1.80(2)	68.97±2.56(5)	70.88±1.53(4)	68.35±2.32(6)	**76.43±1.76(1)**
Eco-11	78.36±0.00(3)	**82.42±1.04(1)**	74.94±2.78(6)	76.86±1.78(5)	79.82±1.88(2)	76.81±1.18(4)
Yea-1	33.09±0.00(6)	59.33±0.60(2)	57.72±1.07(4)	**60.42±0.48(1)**	58.14±0.47(3)	57.30±0.20(5)
Yea-2	18.29±0.00(6)	73.42±0.81(4)	**77.34±0.81(1)**	74.58±0.66(3)	74.98±0.50(2)	69.94±0.32(5)
Yea-3	16.41±0.00(6)	38.79±0.74(4)	33.91±3.07(5)	41.41±1.81(2)	**43.44±2.48(1)**	39.91±0.81(3)
Yea-4	37.48±0.00(5)	47.32±1.08(4)	**52.37±0.51(1)**	48.82±1.15(3)	51.65±0.78(2)	36.71±0.39(6)
Yea-5	22.86±0.00(6)	29.04±2.27(4)	29.61±4.51(2)	29.21±2.35(3)	25.53±4.38(5)	**45.72±2.13(1)**
Yea-6	64.60±0.00(6)	72.74±1.16(3)	72.30±2.51(4)	70.38±1.55(5)	77.00±0.81(2)	**77.25±0.81(1)**
Yea-7	63.81±0.00(5)	21.55±1.72(6)	65.04±0.00(4)	65.14±0.67(2)	65.07±0.65(3)	**65.15±0.67(1)**
Yea-8	**82.82±0.00(1)**	76.29±0.80(6)	82.54±0.89(2)	77.17±0.80(5)	79.32±0.81(3)	78.16±0.62(4)
Yea-9	55.63±0.00(6)	**61.23±0.53(1)**	59.79±1.18(4)	57.91±1.35(3)	59.13±0.75(5)	60.88±0.60(2)
Yea-10	33.13±0.00(5)	39.52±0.82(4)	37.40±1.81(5)	43.52±2.52(2)	**45.72±1.19(1)**	40.91±1.76(3)
Veh-1	40.51±0.00(6)	49.42±0.23(5)	**54.25±0.72(1)**	52.01±1.20(4)	52.04±0.43(3)	52.90±0.73(2)
Veh-2	42.89±0.00(6)	49.01±0.24(5)	55.04±1.46(2)	51.26±1.35(3)	49.84±0.69(4)	**55.29±0.61(1)**
New-1	76.64±0.00(6)	81.92±1.83(5)	94.64±1.51(2)	**97.02±1.00(1)**	85.30±1.51(4)	91.63±1.71(3)
New-2	78.34±0.00(6)	81.09±2.90(5)	93.44±1.52(2)	**95.61±1.46(1)**	82.10±1.28(4)	90.31±1.04(3)
Pima	53.85±0.00(6)	63.63±0.43(2)	**64.46±0.76(1)**	63.20±0.66(4)	62.53±0.34(5)	63.21±0.37(3)
Average rank	4.81	3.15	3.27	3.23	3.38	**3.12**

From Table 10, all the listed classifiers have the similar F1 scores because each of them can outperform others in some specific situations. For instance, the proposed ensemble EBRB decision model has a significant F1 score measure in Yea-5 and Eco-10; E-MLP and E-DT perform much well than others in New-1 and New-2. Although the ensemble EBRB decision model fails to obtain the best F1 score, it does not have the time- consuming training process and has a great interpretability comparing to other classifiers. Thus, the proposed ensemble EBRB decision model can also be treated as more suitable and preferable option to handle imbalanced classification problems.

Table 11. Comparison of average G-mean for ensemble EBRB decision models with conventional ensemble classifiers

Dataset	EBRB	E-SVM	E-DT	E-MLP	E-KNN	Ensemble EBRB
Eco-1	88.63±0.00(3)	89.38±0.18(2)	86.37±0.47(6)	87.77±0.40(5)	**89.98±0.48(1)**	88.49±0.17(4)
Eco-2	92.79±0.00(3)	94.53±0.08(1)	88.41±0.63(6)	91.46±0.34(5)	93.66±0.08(2)	91.85±0.08(4)
Eco-3	71.70±0.00(6)	88.99±0.82(3)	86.57±0.98(5)	89.39±0.23(2)	88.43±0.99(4)	**89.89±0.07(1)**
Eco-4	83.29±0.00(6)	91.23±0.28(3)	90.76±1.83(5)	90.82±0.16(4)	93.24±0.26(2)	**93.75±0.00(1)**
Eco-5	85.22±0.00(5)	**91.16±0.09(1)**	82.58±1.24(6)	88.67±1.83(4)	90.67±0.17(3)	90.83±0.09(2)
Eco-6	82.82±0.00(5)	83.72±1.05(4)	81.78±3.10(6)	88.70±0.91(2)	84.67±0.09(3)	**90.70±0.19(1)**
Eco-7	84.00±0.00(6)	88.98±1.17(3)	85.22±1.80(5)	88.61±1.37(4)	91.44±0.16(1)	90.93±0.10(2)
Eco-8	88.42±0.00(2)	88.00±1.79(3)	**90.60±0.94(1)**	87.11±1.19(6)	87.98±0.93(4)	87.70±0.15(5)
Eco-9	85.01±0.00(6)	**90.59±0.22(1)**	87.22±2.12(5)	87.90±1.78(4)	90.03±0.27(3)	90.13±0.24(2)
Eco-10	78.35±0.00(6)	80.24±1.07(4)	78.48±1.19(5)	81.44±0.93(3)	81.77±1.56(2)	**85.58±1.51(1)**
Eco-11	87.47±0.00(5)	**89.77±0.13(1)**	84.13±1.94(6)	88.37±1.18(3)	89.06±1.55(2)	**87.95±1.04(4)**
Yea-1	45.62±0.00(6)	71.52±0.53(2)	68.74±0.88(5)	**72.45±0.41(1)**	70.56±0.39(3)	69.71±0.16(4)
Yea-2	31.07±0.00(6)	92.24±0.29(2)	87.33±0.71(5)	91.50±0.46(3)	89.54±0.37(4)	**92.50±0.14(1)**
Yea-3	27.59±0.00(6)	79.05±0.58(2)	51.98±2.92(5)	74.19±0.91(3)	70.90±1.58(4)	**81.06±0.46(1)**
Yea-4	51.87±0.00(6)	**87.04±0.66(1)**	68.20±2.80(5)	80.94±0.79(4)	85.08±1.10(3)	86.17±0.75(2)
Yea-5	32.66±0.00(6)	49.11±5.68(3)	38.06±4.88(5)	56.39±4.04(2)	48.17±5.63(4)	**62.25±1.77(1)**
Yea-6	72.94±0.00(6)	**87.80±1.10(1)**	83.06±1.89(5)	85.06±1.07(4)	86.80±1.17(2)	85.27±0.09(3)
Yea-7	72.68±0.00(4)	**74.76±2.58(1)**	72.18±0.00(5)	72.16±1.23(6)	72.75±0.04(3)	72.77±0.05(2)
Yea-8	88.15±0.00(6)	89.42±0.33(4)	88.74±0.69(5)	89.62±0.58(2)	89.57±0.28(3)	**89.95±0.51(1)**
Yea-9	65.93±0.00(6)	**81.07±0.10(1)**	71.90±1.20(5)	78.42±0.69(2)	77.58±0.47(4)	78.02±0.43(3)
Yea-10	46.46±0.00(6)	**68.64±1.39(1)**	52.63±1.17(5)	68.66±2.07(2)	68.54±0.79(3)	67.55±0.71(4)
Veh-1	56.05±0.00(6)	65.42±0.20(5)	67.34±0.67(4)	67.61±1.01(3)	67.73±0.39(2)	**68.32±0.64(1)**
Veh-2	57.92±0.00(6)	65.14±0.21(5)	67.68±1.21(2)	66.86±1.27(3)	66.08±0.62(4)	**70.57±0.50(1)**
New-1	78.99±0.00(6)	89.16±0.96(5)	97.92±1.14(2)	**98.90±0.69(1)**	92.68±1.07(4)	95.11±1.26(3)
New-2	80.53±0.00(6)	90.66±0.89(5)	96.73±0.82(2)	**98.11±1.41(1)**	90.69±0.85(4)	96.06±1.12(3)
Pima	63.29±0.00(6)	71.67±0.37(2)	**72.17±0.61(1)**	71.04±0.56(4)	70.69±0.28(5)	71.27±0.31(3)
Average rank	5.42	2.54	4.50	3.19	3.04	**2.31**

From Table 11, the proposed ensemble EBRB decision model obtains the best G-mean and it is significantly better than the other listed classifiers. This is due to the enhancement of the diversity of each base EBRB constructed by the diverse data oversampling methods, the calculation of attribute weights based on information gain, and the combination process based on the consistency of activated rules.

Preliminary conclusions are summarized based on the comparative results:

(1) The proposed ensemble EBRB decision model obtains the best F1 score. Although the winning margin is not large, an overwhelming performance was provided by the ensemble EBRB decision model on some of the datasets, *e.g.*, Yea-5 and Eco-10, comparing to other machine learning algorithms. For G-mean, the ensemble EBRB decision model outperforms other classifiers. Hence, it is believed that the proposed model can handle imbalanced classification problems properly.

(2) It is impossible to find the classifier that can produce the best performance for all datasets because many factors, such as structure, noise, and the size of training data may impact the performance of the classifier. Thus, decision-makers should make choices according to the different situations.

In summary, the proposed ensemble EBRB decision model provides a satisfying evaluation result in F1 score and outperforms other machine learning classifiers in G-mean with a large winning margin. Hence, for a specific imbalanced classification scenario, the ensemble EBRB decision model can be considered as an option to give the accurate predictions.

6. Conclusions

In this study, a novel ensemble classifier based on EBRB decision model, called ensemble EBRB decision model, was proposed for imbalanced classification problems, where the ensemble EBRB decision model consists of two components: diversity-based base EBRB construction and consistency-based ensemble EBRB inference. 26 imbalanced classification datasets were used to validate the effectiveness of the proposed ensemble EBRB decision model by comparing with some popular classifiers. The main conclusions of this study can be summarized into the following three aspects:

(1) A diversity-based base EBRB construction method was proposed to improve the conventional EBRB construction scheme in the aim of increasing the diversity of the ensemble EBRB decision model, because the diversity of components is of great importance but hard to be achieved in an ensemble classifier. In the proposed construction method, the diversity of base EBRBs is enhanced by using random sampling and oversampling techniques to rebalance the original imbalanced dataset, as well as the use of information gain to calculate the attribute weights of each base EBRBs.

(2) A consistency-based ensemble EBRB inference method was proposed to improve the conventional EBRB inference scheme in the aim of increasing the consistency of activated rules, because the class overlap is the direct cause of the activated rules' inconsistency and it tends to be more serious in the context of class imbalance. In the proposed inference method, the integrated belief distributions are used together with the consistency of activated rules to produce a final belief distribution, so that the classification performance of the ensemble EBRB decision model can be improved.

(3) Experimental comparisons on imbalanced datasets demonstrated that the proposed ensemble EBRB decision model has better performance than other EBRB decision models. Compared to traditional machine learning algorithm-based ensemble classifiers, the proposed model showed promising performance in F1 score and outperformed the others with a large winning margin in G-mean. From those experimental comparisons, the ensemble EBRB decision model could be a competent option to tackle imbalanced classification problems.

For future researches, the number of reference values for each antecedent attribute and their corresponding utility values is worth further in-depth study, because they are important elements to construct an effective and efficient EBRB but failed to be considered in the proposed diversity-based base EBRB construction method. Additionally, future research can focus on how to adaptively set activation factor for determining the suitable activated rules when an EBRB decision model is applied to handle practical problems with imbalanced data.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Nos. 72001043 and 61773123), the Natural Science Foundation of Fujian Province, China (No. 2020J05122), the Humanities and Social Science Foundation of the Ministry of Education, China (No. 20YJC630188), and the Social Science Foundation of Fujian Province, China (No. FJ2019C032).

Reference

[1] J. Li, S. Fong, R.K. Wong, V.W. Chu, Adaptive multi-objective swarm fusion for imbalanced data classification, *Information Fusion*, 2018, 39: 1-24.

[2] D. Gan, J. Shen, B. An, M. Xu, N. Liu, Integrating TANBN with cost sensitive classification algorithm for imbalanced data in medical diagnosis, *Computers & Industrial Engineering*, 2020, 140: 106266.

[3] Z. Li, M. Huang, G. Liu, C. Jiang, A hybrid method with dynamic weighted entropy for handling the problem of class imbalance with overlap in credit card fraud detection, *Expert Systems with Applications*, 2021, 175: 114750.

[4] H.B. He, Y. Bai, E.A. Garcia, S.T. Li, ADASYN: Adaptive synthetic sampling approach for imbalanced learning, *IEEE International Joint Conference on Neural Networks*, 2008, 1322-1328.

[5] C. Bunkhumpornpat, K. Sinapiromsaran, C. Lursinsap, Safe-Level-SMOTE: Safe-Level-Synthetic Minority Over-Sampling TEchnique for Handling the Class Imbalanced Problem, *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2009, 5476: 475-482.

[6] H. Han, W.Y. Wang, B.H. Mao, Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning, *Advances in Intelligent Computing*, 2005: 878-887.

[7] J. Liu, L. Martinez, A. Calzada, H. Wang, A novel belief rule base representation, generation and its inference methodology, *Knowledge-Based Systems,* 2013, 53: 129-141.

[8] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: Synthetic Minority Over-sampling Technique, *Journal of Artificial Intelligence Research*, 2012, 16: 321-357

[9] J.F. Díez-Pastor, J.J. Rodríguez, C.I. García-Osorio, L.I. Kuncheva, Diversity techniques improve the performance of the best imbalance learning ensembles, *Information Sciences*, 2015, 325: 98-117.

[10] L.H. Yang, J. Liu, Y.M. Wang, C. Nugent, L. Martínez, Online updating extended belief rule-based system for sensor-based activity recognition, *Expert Systems with Applications*, 2021, 186: 115737.

[11] F.F. Ye, S. Wang, P. Nicholl, L.H. Yang, Y.M. Wang, Extended belief rule-based model for environmental investment prediction with indicator ensemble selection, *International Journal of Approximate Reasoning*, 2020, 126: 290-307.

[12] L.H. Yang, J. Liu, Y.M. Wang, L. Martínez, A Micro-Extended Belief Rule-Based System for Big Data Multiclass Classification Problems, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, 51(1): 420-440.

[13] L.H. Yang, Y.M. Wang, Y.X. Lan, L. Chen, Y.G. Fu, A data envelopment analysis (DEA)-based method for rule reduction in extended belief-rule-based systems, *Knowledge-Based Systems*, 2017, 123: 174-187.

[14] A. Zhang, F. Gao, M. Yang, W. Bi, A new rule reduction and training method for extended belief rule base based on DBSCAN algorithm, *International Journal of Approximate Reasoning*, 2020, 119: 20-39.

[15] H.Z. Zhu, M.Q. Xiao, X. Zhao, X.L. Tang, L.H. Yang, W.J. Kang, Z.Z. Liu, A structure optimization method for extended belief-rule-based classification system, *Knowledge-Based Systems*, 2020, 203: 106096.

[16] A. Calzada, J. Liu, H. Wang, A. Kashyap, A New Dynamic Rule Activation Method for Extended Belief Rule-Based Systems, *IEEE Transactions on Knowledge and Data Engineering*, 2015, 27(4):880-894.

[17] T.Y. Ren, F.F. Ye, L.H. Yang, J. Liu, Y.Y. Wang, Dynamic Rule Activtion Method Based on Activation Factor for Extended Belief Rule-based Systems, *International Conference on Intelligent Systems and Knowledge Engineering*, 2021, In Press.

[18] L.H. Yang, Y.M. Wang, Y.G. Fu, A consistency analysis-based rule activation method for extended belief-rule-based systems, *Information Sciences*, 2018, 445-446: 50-65.

[19] L.H. Yang, Y.M. Wang, Q. Su, Y.G. Fu, K.S. Chin, Multi-attribute search framework for optimizing extended belief rule-based systems, *Information Sciences*, 2016, 370-371: 159-183.

[20] Y.G. Fu, J.H. Zhuang, Y.P. Chen, L.K. Guo, Y.M. Wang, A framework for optimizing extended belief rule base systems with improved Ball trees, *Knowledge-Based Systems*, 2020, 210: 106484.

[21] D.L. Xu, J. Liu, J.B. Yang, G.P. Liu, J. Wang, L. Jenkinson, J. Ren, Inference and learning methodology of belief-rule-based expert system for pipeline leak detection, *Expert Systems with Applications*, 2007, 32(1): 103-113.

[22] Y.M. Wang, J.B. Yang, D.L. Xu, Environmental impact assessment using the evidential reasoning approach, *European Journal of Operational Research*, 2006, 174(3): 1885-1913.

[23] L.H. Yang, J. Liu, Y.M. Wang, H. Wang, L. Martínez, Enhancing extended belief rule-based systems for classification problems using decomposition strategy and overlap function, *International Journal of Machine Learning and*

Cybernetics, 2021, In Press.

[24] J. Alcala-Fdez, A. Fernandez, J. Luengo, J. Derrac, S. Garcia, L. Sanchea, F. Herrera, KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework, *Journal of Multiple- Valued Logic and Soft Computing*, 2010, 17: 255-287.

[25] L.H. Yang, J. Liu, Y.M. Wang, L. Martínez, Extended belief-rule-based system with new activation rule determination and weight calculation for classification problems, *Applied Soft Computing*, 2018, 72: 261-272.

[26] H.X. Guo, Y.J. Li, J. Shang, M.Y. Gu, Y.Y. Huang, B. Gong, Learning from class-imbalanced data: Review of methods and applications, *Expert Systems with Applications*, 2017, 73: 220-239.

[27] W.C. Lin, C.F. Tsai, Y.H. Hu, J.S. Jhang, Clustering-based undersampling in class-imbalanced data, *Information Sciences*, 2017, 409: 17-26.

[28] H.L. Yu, C.X. Mu, C.Y. Sun, W.K. Yang, X.B. Yang, X. Zuo, Support vector machine-based optimized decision threshold adjustment strategy for classifying imbalanced data, *Knowledge-Based Systems*, 2015, 76: 67-78.

[29] C.G. Zhang, W. Gao, J.Z. Song, J.Q. Jiang, An imbalanced data classification algorithm of improved autoencoder neural network, *Eighth International Conference on Advanced Computational Intelligence*, 2016: 85-99.

[30] F.Y. Cheng, J. Zhang, C.H. Wen, Cost-sensitive large margin distribution machine for classification of imbalanced data, *Pattern Recognition Letters*, 2016, 80: 107-112.

[31] Y. Zhang, P.P. Fu, W.Z. Liu, G.L. Chen, Imbalanced data classification based on scaling kernel-based support vector machine, *Neural Computing and Applications*, 2014, 25(3): 927-935.

[32] S. Kim, H. Kim, Y. Namkoong, Ordinal classification of imbalanced data with application in emergency and disaster information services, *IEEE Intelligent Systems*, 2016, 31(5): 50-56.

[33] L.X. Duan, M.Y. Xie, T.B. Bai, J.J. Wang, Support vector data description for machinery multi-fault classification with unbalanced datasets, *IEEE International Conference on Prognostics and Health Management*, 2016: 1-6.

[34] Y.M. Sun, M.S. Kamel, A.K. Wong, Y. Wang, Cost-sensitive boosting for classification of imbalanced data, *Pattern Recognition*, 2007, 40(12): 3358-3378.

[35] S. Ali, A. Majid, S.G. Javed, M. Sattar, Can-CSC-GBE: Developing Cost-sensitive Classifier with Gentleboost Ensemble for breast cancer classification using protein amino acids and imbalanced data, *Computers in Biology and Medicine*, 2016, 73: 38-46.

[36] X. Wan, J.M. Liu, W.K. Cheung, T.J. Tong, Learning to improve medical decision making from imbalanced data without a priori cost. *BMC Medical Informatics and Decision Making*, 2014, 14(1): 1-9.

[37] R. Barandela, R.M. Valdovinos, J.S. Sánchez, New applications of ensembles of classifiers, *Pattern Analysis & Applications*, 2003, 6(3): 245-256.

[38] X.Y. Liu, J.X. Xu, Z.H. Zhou, Exploratory undersampling for class-imbalance learning, *IEEE Transactions on Systems, Man, and Cybernetics, Part B* (*Cybernetics*), 2018, 39(2): 539-550.

[39] S. Wang, X. Yao, Diversity analysis on imbalanced data sets by using ensemble models, *IEEE Symposium on Computational Intelligence and Data Mining*, 2009: 324-331.

[40] N.V. Chawla, A. Lazarevic, L.O. Hall, K.W. Bowyer, SMOTEBoost: Improving prediction of the minority class in boosting, *European Conference on Principles of Data Mining and Knowledge Discovery*, 2003: 107-119.

[41] P. Lim, C.K. Goh, K.C. Tan, Evolutionary cluster-based synthetic oversampling ensemble (eco-ensemble) for imbalance learning, *IEEE Transactions on Cybernetics*, 2016, 47(9): 2850-2861.

[42] X. Gao, B. Ren, H. Zhang, B.H. Sun, J.L. Li, J.H. Xu, Y. He, K.S. Li, An ensemble imbalanced classification method based on model dynamic selection driven by data partition hybrid sampling, *Expert Systems with Applications*, 2020, 160: 113660.

[43] Y.G. Fu, H.Y. Huang, Y. Guan, Y.M. Wang, W.X. Liu, W.J. Fang, EBRB cascade classifier for imbalanced data via rule weight updating, *Knowledge-Based Systems*, 2021, 223: 107010.

[44] M. Juez-Gil, A. Arnaiz-Gonzalez, J.J. Rodriguez, C. Lopez-Nozai, C. Garcia-Osorio, Approx-SMOTE: Fast SMOTE for Big Data on Apache Spark, *Neurocomputing*, 2021, 464: 432-437.

[45] R. Kohavi, D.H. Wolpert, Bias plus variance decomposition for zero-one loss functions, In *Proceeding of the 13rd International Conference on Machine Learning*, 1996: 275-283.

[46] L.I. Kuncheva, C.J. Whitaker, Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy, *Machine Learning*, 2003, 51(2): 181-207.

[47] Z. Chen, J. Duan, L. Kang, G. Qiu, A hybrid data-level ensemble to enable learning from highly imbalanced dataset, *Information Sciences,* 2021, 554: 157-176.

[48] A. Anand, G. Pugalenthi, G.B. Fogel, P.N. Suganthan, An approach for classification of highly imbalanced data using weighting and undersampling, *Amino acids*, 2010, 39(5): 1385-1391.