



**HAL**  
open science

# Reliability-Based Imbalanced Data Classification with Dempster-Shafer Theory

Hongpeng Tian, Zuowei Zhang, Arnaud Martin, Zhunga Liu

► **To cite this version:**

Hongpeng Tian, Zuowei Zhang, Arnaud Martin, Zhunga Liu. Reliability-Based Imbalanced Data Classification with Dempster-Shafer Theory. 7th International Conference on Belief Functions, Oct 2022, PARIS, France. 10.1007/978-3-031-17801-6\_8 . hal-03816207

**HAL Id: hal-03816207**

**<https://hal.science/hal-03816207>**

Submitted on 15 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Metadata of the chapter that will be visualized in SpringerLink

Book Title	Belief Functions: Theory and Applications	
Series Title		
Chapter Title	Reliability-Based Imbalanced Data Classification with Dempster-Shafer Theory	
Copyright Year	2022	
Copyright HolderName	The Author(s), under exclusive license to Springer Nature Switzerland AG	
Author	Family Name	<b>Tian</b>
	Particle	
	Given Name	<b>Hongpeng</b>
	Prefix	
	Suffix	
	Role	
	Division	School of Electrical Engineering
	Organization	Zhengzhou University
	Address	Zhengzhou, China
	Email	tianhongpeng1220@163.com
Corresponding Author	Family Name	<b>Zhang</b>
	Particle	
	Given Name	<b>Zuowei</b>
	Prefix	
	Suffix	
	Role	
	Division	School of Automation
	Organization	Northwestern Polytechnical University
	Address	Xi'an, China
	Division	
	Organization	Univ Rennes, CNRS, IRISA
	Address	Rue E. Branly, 22300, Lannion, France
	Email	zuowei_zhang@mail.nwpu.edu.cn
Author	Family Name	<b>Martin</b>
	Particle	
	Given Name	<b>Arnaud</b>
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Univ Rennes, CNRS, IRISA
	Address	Rue E. Branly, 22300, Lannion, France
	Email	Arnaud.Martin@univ-rennes1.fr
Author	Family Name	<b>Liu</b>
	Particle	

Given Name	<b>Zhunga</b>
Prefix	
Suffix	
Role	
Division	School of Automation
Organization	Northwestern Polytechnical University
Address	Xi'an, China
Email	liuzhunga@nwpu.edu.cn

---

**Abstract**

The classification analysis of imbalanced data remains a challenging task since the base classifier usually focuses on the majority class and ignores the minority class. This paper proposes a reliability-based imbalanced data classification approach (RIC) with Dempster-Shafer theory to address this issue. First, based on the minority class, multiple under-sampling for the majority one are implemented to obtain the corresponding balanced training sets, which results in multiple globally optimal trained classifiers. Then, the neighbors are employed to evaluate the local reliability of different classifiers in classifying each test sample, making each global optimal classifier focus on the sample locally. Finally, the revised classification results based on various local reliability are fused by the Dempster-Shafer (DS) fusion rule. Doing so, the test sample can be directly classified if more than one classifier has high local reliability. Otherwise, the neighbors belonging to different classes are employed again as the additional knowledge to revise the fusion result. The effectiveness has been verified on synthetic and several real imbalanced datasets by comparison with other related approaches.

---

**Keywords**  
(separated by '-')

Imbalanced data - Reliability - Dempster-Shafer theory

---



# Reliability-Based Imbalanced Data Classification with Dempster-Shafer Theory

Hongpeng Tian<sup>1</sup>, Zuowei Zhang<sup>2,3</sup>(✉), Arnaud Martin<sup>3</sup>, and Zhunga Liu<sup>2</sup>

<sup>1</sup> School of Electrical Engineering, Zhengzhou University, Zhengzhou, China

<sup>2</sup> School of Automation, Northwestern Polytechnical University, Xi'an, China  
zuowei\_zhang@mail.nwpu.edu.cn, liuzhunga@nwpu.edu.cn

<sup>3</sup> Univ Rennes, CNRS, IRISA, Rue E. Branly, 22300 Lannion, France  
Arnaud.Martin@univ-rennes1.fr

**Abstract.** The classification analysis of imbalanced data remains a challenging task since the base classifier usually focuses on the majority class and ignores the minority class. This paper proposes a reliability-based imbalanced data classification approach (RIC) with Dempster-Shafer theory to address this issue. First, based on the minority class, multiple under-sampling for the majority one are implemented to obtain the corresponding balanced training sets, which results in multiple globally optimal trained classifiers. Then, the neighbors are employed to evaluate the local reliability of different classifiers in classifying each test sample, making each global optimal classifier focus on the sample locally. Finally, the revised classification results based on various local reliability are fused by the Dempster-Shafer (DS) fusion rule. Doing so, the test sample can be directly classified if more than one classifier has high local reliability. Otherwise, the neighbors belonging to different classes are employed again as the additional knowledge to revise the fusion result. The effectiveness has been verified on synthetic and several real imbalanced datasets by comparison with other related approaches.

[AQ1](#)

**Keywords:** Imbalanced data · Reliability · Dempster-Shafer theory

## 1 Introduction

Imbalanced data refers to the dataset has an unequal distribution between classes [1]. For a binary class problem, if the number of samples in the majority class is significant larger than that of the minority class, traditional classifiers, such as  $K$ -nearest neighbors (K-NN) [2], support vector machine classifier (SVM) [3], are dedicated to maximize the overall classification performance. In this case, most minority samples are assigned to majority class.

Increasingly works are emerged for classifying imbalanced data, and they can be roughly divided into three categories including sampling approaches [4], cost-sensitive learning [5] and ensemble learning [6]. Sampling approaches focus

on preprocessing the input data to balance the classes. By doing this, the preprocessed data can be classified by basic classifiers. Cost-sensitive learning approaches assign relatively high weights to minority samples, which can reduce the misclassification of the minority class. Ensemble learning approaches combine different classifiers trained by various subsets, which supplies the complementarity information to improve the performance of classification with respect to an individual classifier. However, these imbalanced data classification approaches only consider the global optimum and are not suitable for each test sample. For instance, the samples lying in the overlap area of different classes are indistinguishable and easily misclassified. In this case, there are some uncertain information between classes.

Dempster-Shafer theory (DST) [7, 8], also known as the theory of belief functions, has the advantage of reasoning uncertain information, and has been widely used in classification [9–12]. Recently, a few works [13, 14] have been proposed to deal with imbalanced data classification within the belief function theory. Although these approaches has the advantage of capturing uncertain information thanks to evidence reasoning, they fail to consider the local performance of classifiers for each test sample. In this paper, we propose a reliability-based imbalanced data classification approach with Dempster-Shafer theory. The contributions mainly include three aspects. 1) We design a reliability evaluation strategy to obtain local reliability of different classifiers for each test sample, which can characterize the local performance of classifiers. 2) We introduce a revision strategy to resubmit the samples with low local reliability of different classifiers according to neighbors from various classes. 3) We apply RIC to synthetic and several real imbalanced datasets to demonstrate the superiority.

The rest of this paper is organized as follows. The proposed approach is presented in detail in Sect. 2. Then, it is tested in Sect. 3 and compared with several other typical methods, followed by conclusions.

## 2 Reliability-Based Imbalanced Data Classification

In this section, a reliability-based imbalanced data classification approach is proposed in detail. Assume that a test set  $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$  is classified under the frame of discernment  $\Omega = \{\omega_{min}, \omega_{maj}\}$  according to a training set  $\mathcal{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_M\}$  on  $H$  different attribute spaces.  $\mathcal{Y}_{min}$  and  $\mathcal{Y}_{maj}$  represent the minority class and majority class, respectively.

### 2.1 Multiple Under-Sampling for Majority Class

In this subsection, we implement random under-sampling<sup>1</sup> for the majority class multiple times to obtain different training sets thereby training basic classifiers.

$T$  subsets  $\mathcal{Y}_{maj}^1, \dots, \mathcal{Y}_{maj}^T$  are random sampled from the majority class  $\mathcal{Y}_{maj}$ . Each subset has the same number of samples as that of the minority class  $\mathcal{Y}_{min}$ ,

<sup>1</sup> In applications, users can employ other appropriate under-sampling approaches according to the request of practice.

and is combined with  $\mathcal{Y}_{min}$  to form a new training set. By doing this, we can obtain  $T$  training sets, named  $\mathcal{Y}^1, \dots, \mathcal{Y}^T$ , and the number of  $T$  is denoted as:

$$T = [IR] \quad (1)$$

with

$$IR = \frac{|Y_{maj}|}{|Y_{min}|} \quad (2)$$

where  $IR$ , such that  $IR \geq 1$ , refers to the measurement for the imbalanced degree of the dataset, and  $|\cdot|$  represents the cardinality symbol.  $[\cdot]$  is a rounding symbol that rounds the elements of  $IR$  to the nearest integers towards infinity.

Then, each training set can train a basic classifier that has high performance in classifying a balanced dataset. The classification result of  $\mathbf{x}_i$  by  $t$ -th classifier is denoted as  $\mathcal{P}_i^t = [p_i^t(\{\omega_{min}\}), p_i^t(\{\omega_{maj}\})]$ ,  $t = 1, \dots, T$ .

## 2.2 Evaluate the Local Reliability for Classifiers Fusion

In this subsection, we evaluate the local reliability of different classifiers for classifying each test sample. Then, we combine classifiers with various reliability by the original discounting fusion rule.

Here, we employ the neighbors  $\mathbf{y}_1, \dots, \mathbf{y}_K$  of the test sample  $\mathbf{x}_i$  to evaluate the local reliability of different classifiers, since  $\mathbf{x}_i$  has the similar data structure and distribution with respect to  $\mathbf{y}_1, \dots, \mathbf{y}_K$ . The better the performance of the classifier to classify  $\mathbf{y}_1, \dots, \mathbf{y}_K$ , the higher the reliability for classifying  $\mathbf{x}_i$ . Based on the above analysis, we define a rule to evaluate the degree of reliability of different classifiers, denoted as:

$$\xi_{it} = \frac{\exp(-\vartheta_{it})}{\sum_{t=1}^T \exp(-\vartheta_{it})} \quad (3)$$

with

$$\vartheta_{it} = \sum_{k=1}^K \sqrt{\sum_{\{\omega_c\} \in \Omega} [p_k^t(\{\omega_c\}) - l_k(\{\omega_c\})]^2} \quad (4)$$

where  $\xi_{it}$ , such that  $0 < \xi_{it} < 1$ , represents the reliability of the  $t$ -th classifier for classifying  $\mathbf{x}_i$ .  $p_k^t(\{\omega_c\})$  refers to the probability of  $\mathbf{y}_k$  belongs to  $\{\omega_c\}$ . The truth of classification of  $\mathbf{y}_k$  is characterized by the binary vector  $L_k = [l_k(\{\omega_{min}\}), l_k(\{\omega_{maj}\})]$ . The  $l_k(\{\omega_c\}) = 1$  if the true class of  $\mathbf{y}_k$  is  $\{\omega_c\}$ . If not,  $l_k(\{\omega_c\})$  is equal to 0.  $\Omega$  is the frame of discernment, such that  $\Omega = \{\{\omega_{min}\}, \{\omega_{maj}\}\}$ . We can observe that the lower the deviation between classification results and truths, the higher reliability of the classifier.

Each classification result can be considered as a piece of evidence under the framework of DST, which is appealing to combine multi-source information. The reliability-based discount fusion method [8], is employed here for discounting and fusing pieces of evidence. The reliability  $\xi_{it}$  for different  $T$  classifiers can be considered as the discounting factors. The discounted masses of belief is denoted as:

$$\begin{cases} m_i^t(\{\omega_c\}) = \xi_{it} p_i^t(\{\omega_c\}), \{\omega_c\} \in \Omega \\ m_i^t(\Omega) = 1 - \xi_{it} \end{cases} \quad (5)$$

where  $p_i^t(\{\omega_c\})$  represents the probability (Bayesian BBA) that the sample  $\mathbf{x}_i$  belongs to  $\{\omega_c\}$  under Bayesian framework, and  $\Omega$  is the the total unknown class. We can find that the more important and reliable the classification result, the larger the corresponding discounting factor, and the less discounted information assigned to the total ignorance  $\Omega$ . In particular, the degree of conflict between pieces of evidence is reduced since the conflict information is transferred into  $\Omega$  that plays a particular neutral role in the fusion process. In this case, the global fusion result for  $T$  basic belief assignments (BBAs) of the sample  $\mathbf{x}_i$  is denoted as:

$$\mathbf{m}_i = \mathbf{m}_i^1 \oplus \cdots \oplus \mathbf{m}_i^T \quad (6)$$

where  $\mathbf{m}_i^t = [m_i^t(\{\omega_{min}\}), m_i^t(\{\omega_{maj}\})]$ , and  $\oplus$  refers to the DS fusion rule [7] for the combination of these  $T$  pieces of evidence.  $\mathbf{m}_i$  represents the normalized combination result. As a result, the fused BBAs can be transferred into pignistic probability[15] for the preliminary decision-making.

### 2.3 Employ Neighbors for Final Decision

In this subsection, we employ neighbors from different classes as the additional information to make final decision.

For the test sample  $\mathbf{x}_i$ , it can be directly classified if there is more than one classifier that has high local reliability. In contrast, when all classifiers have low reliability to classify  $\mathbf{x}_i$ , which means it is hard to be correctly classified by different classifiers.

We evaluate the different degrees of local reliability of classifiers for classifying  $\mathbf{x}_i$  before normalization, and obtain the max values of them, denoted as:

$$\widehat{\xi}_{i,\max} = \max\{\widehat{\xi}_{i1}, \dots, \widehat{\xi}_{iT}\} \quad (7)$$

where  $\widehat{\xi}_{it}$  represents the degree of local reliability of  $t$ -th classifier for classifying  $\mathbf{x}_i$ , such that  $\widehat{\xi}_{it} = \exp(-\vartheta_{it})$ . The higher the value of  $\widehat{\xi}_{i,\max}$ , the bigger the possibility of  $\mathbf{x}_i$  being correctly classified. Thus, we define a threshold  $\delta$  to distinguish whether  $\mathbf{x}_i$  can be directly classified or not, given by:

$$\delta = \text{quantile}(\widehat{\Xi}_{\max}, \gamma) \quad (8)$$

with

$$\widehat{\Xi}_{\max} = \{\widehat{\xi}_{1,\max}, \dots, \widehat{\xi}_{N,\max}\} \quad (9)$$

where  $\gamma$  is a quantile number such that  $\gamma \in [0, 1]$ . If  $\widehat{\xi}_{i,\max} > \delta$ , the test sample  $\mathbf{x}_i$  can be classified directly according to the classification result obtained by discounting fusion. Otherwise, we need to mine some additional information by neighbors to revise classification results. We convert the distances between  $\mathbf{x}_i$

and different classes into the mass of belief that it belongs to different classes, denoted as:

$$\hat{m}_i(\{\omega_c\}) = \frac{\exp(-d(\mathbf{x}_i, \{\omega_c\}))}{\sum_{\{\omega_c\} \in \Omega} \exp(-d(\mathbf{x}_i, \{\omega_c\}))}. \quad (10)$$

where  $d(\mathbf{x}_i, \{\omega_c\})$  represent the mean Euclidean distance between  $\mathbf{x}_i$  and its  $K$  neighbors in class  $\{\omega_c\}$ . We can observe that the larger the distance  $d(\mathbf{x}_i, \{\omega_c\})$ , the lower the possibility that  $\mathbf{x}_i$  belongs to class  $\{\omega_c\}$ . Then, the BBA of  $\mathbf{x}_i$ , named  $\hat{\mathbf{m}}_i$ , is fused with  $\mathbf{m}_i$  according to the DS fusion rule. Finally, the fused BBAs can be transferred into pignistic probability to make final decision.

### 3 Experiment Applications

In this section, the proposed RIC is compared with several typical approaches including ROS [4], RUS [16], SMOTE [17], CBU [18] and RUSBoost [6]. SVM [3] is taken as the basic classifier in different approaches. Two common indexes [1], *i.e.*, F-measure (FM) and G-mean (GM), widely used in imbalanced data classification, are employed to evaluate the performance of different approaches. The higher the values of FM and GM, the better the performance of the approach.

#### 3.1 Benchmark Datasets

A 2-D dataset with two classes  $\Omega = \{\omega_{min}, \omega_{maj}\}$  is given in Fig. 1(a)(b), where each sample denoted as a point has two dimensions of attributes corresponding to  $x$ -coordinate and  $y$ -coordinate. The minority class  $\omega_{min}$  has 2000 samples and majority class  $\omega_{maj}$  consists of 200 samples. All the samples are generated from two bivariate Gaussian densities and have the following means vectors and covariance matrices, denoted as:  $\mu_{min} = (3.1, 5)$ ,  $\Sigma_{min} = 0.01\mathbf{I}$ ,  $\mu_{maj} = (4, 5)$ ,  $\Sigma_{maj} = 0.1\mathbf{I}$ , where  $\mathbf{I}$  represents the  $2 \times 2$  identity matrix. Half of the samples in each class are randomly selected as training samples and others are as test samples.  $\omega_{min}^{tr}$  and  $\omega_{maj}^{te}$  represent the minority class and majority class in the training set, respectively. The ground truth of test set is marked by different colors and represented by  $\omega_{min}^{te}$  and  $\omega_{maj}^{te}$ .

Ten generally used real imbalanced datasets from Keel repository<sup>2</sup> are employed to test and evaluate the performance of different approaches in classifying imbalanced data. Each dataset is partitioned using a five-folds stratified cross validation. The basic information of these datasets including the number of all samples (#Size.), majority class samples (#Maj.), minority class samples (#Min), attributes (#Attr.) and imbalance ratio (#IR.) are shown in Table 1.

<sup>2</sup> <http://www.keel.es/>.



**Table 1.** Basic information of the Keel datasets.

Data	#Size.	#Min.	#Maj.	#Attr.	#IR.
glass1	214	76	138	9	1.82
yeast1	1484	429	1055	8	2.46
vehicle2	846	218	628	18	2.88
ecoli2	336	52	284	7	5.46
page-blocks0	5472	559	4913	10	8.79
vowel0	998	90	898	13	9.98
led7digit	443	37	406	7	10.97
ecoli4	336	20	316	7	15.80
yeast5	1484	44	1440	8	32.73
shuttle2vs5	3316	49	3267	9	66.67

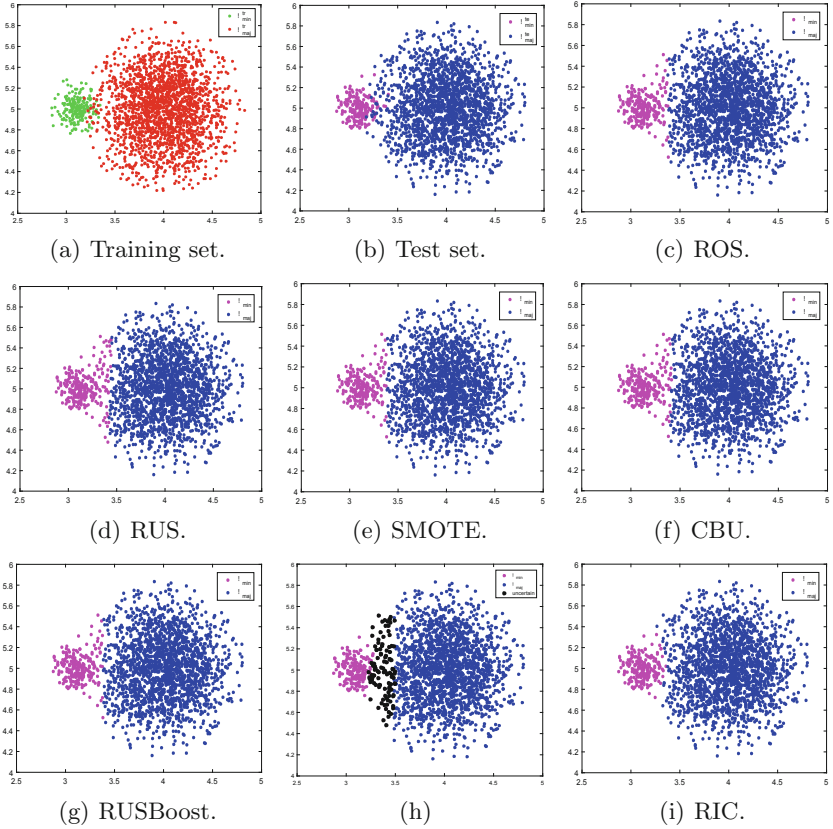
### 3.2 Performance Evaluation

#### 1) RIC vs. comparisons in the synthetic dataset.

This experiment is designed to intuitively validate the effectiveness of RIC in classifying imbalanced data with the overlapped area between classes. We take  $K = 7$  and  $\gamma = 0.05$  in RIC, and the parameters in comparison approaches are as default. The training and test set of the synthetic dataset are reported in Fig. 1(a)(b). We can see that the majority class and minority class are partly overlapped on their borders. The classification results of comparison approaches are shown in Fig. 1(c)-(g), where most samples lying in the overlapped area are misclassified. Actually, these samples are hard to be correctly classified and there are uncertain information between classes in this case. We can observe from Fig. 1(h) that these samples marked by black points are correctly identified. As shown in Fig 1(i), RIC can correctly resubmit most of these samples according to neighbors from different classes. Moreover, an ablation study is carried out on this dataset to compare the contribution of each step in RIC, and the results are reported in Table 2. We can find that with the addition of each step in RIC, the performance continues to improve, which verifies the effectiveness of each step in RIC.

#### 2) RIC vs. comparisons in Keel datasets.

In this experiment, ten Keel datasets are employed to further investigate the effectiveness of the proposed approach by comparing it to other comparison approaches in real word datasets. We take  $K = 7$  and  $\gamma = 0.15$  in RIC, and the parameters in comparison approaches are as default. The classification results of different approaches for classifying different approaches are reported in Table 3. We can observe that the proposed RIC generally provides better performances than comparison approaches in most datasets. The reason is that RIC evaluates the local reliability of different classifiers in classifying each test sample, making each global optimal classifier focus on the sample locally.



**Fig. 1.** Classification results of the synthetic dataset by different approaches.

**Table 2.** The values of FM and GM of different approaches for the synthetic dataset (IN %)

Indexes	ROS	RUS	SMOTE	CBU	RUSBoost	RIC <sub>Step 1</sub>	RIC <sub>Step 1-2</sub>	RIC <sub>Step 1-3</sub>
FM	90.45	85.84	90.25	88.30	90.25	85.47	86.02	<b>94.06</b>
GM	98.72	98.34	98.70	98.67	98.70	98.29	98.36	<b>98.92</b>

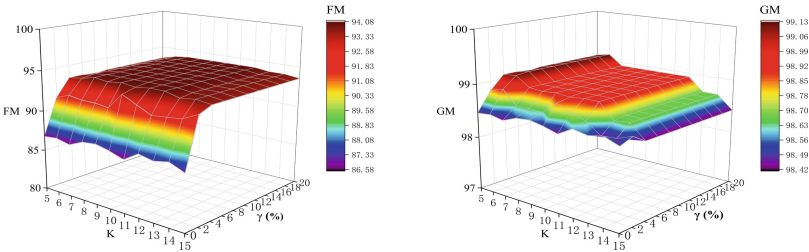
### 3.3 Influence of $K$ and $\gamma$

In this experiment, we employ the synthetic dataset to investigate the performance of RIC with various values of  $K$  and  $\gamma$ . The classification results of RIC with various parameters are shown in Fig. 2, where the  $x$ -coordinate denotes the value of  $K$ , ranging from 5 to 15, and the  $y$ -coordinate represents the value of  $\gamma$ , which is expressed in  $[0, 1]$ . The  $z$ -coordinate of Fig. 2(a) and (b) is the value of FM and GM, respectively. We can see that with the increase of  $K$ , the variations of the result are small, which verifies it is robust to the value of  $K$ . Of course,

**Table 3.** Classification results for Keel datasets by different approaches (IN %)

Indexes	Datasets	ROS	RUS	SMOTE	CBU	RUSBoost	RIC
	ecoli2	72.38	72.40	71.45	70.97	69.30	<b>83.70</b>
	ecoli4	73.21	60.64	74.35	63.69	68.55	<b>84.14</b>
	led7digit	61.99	58.16	64.44	60.15	60.16	<b>80.03</b>
	glass1	57.36	56.59	55.81	56.10	49.76	<b>66.26</b>
FM	page-blocks0	22.87	22.59	18.38	31.54	46.84	<b>80.54</b>
	shuttle2vs5	93.61	74.01	93.61	73.64	93.02	<b>97.89</b>
	vehicle2	92.04	91.90	92.79	89.42	80.80	<b>94.03</b>
	vowel0	78.23	72.45	81.58	72.29	75.25	<b>94.10</b>
	yeast1	58.15	58.29	58.57	57.24	57.87	<b>58.84</b>
	yeast5	46.69	38.65	46.96	37.78	56.96	<b>58.41</b>
	ecoli2	90.77	89.72	90.39	89.15	88.05	<b>93.36</b>
	ecoli4	95.28	95.08	92.92	93.79	87.00	<b>96.20</b>
	led7digit	88.15	89.66	87.51	88.82	87.95	<b>90.04</b>
	glass1	61.11	54.14	54.63	56.55	64.40	<b>71.87</b>
GM	page-blocks0	50.26	46.02	38.21	48.85	83.51	<b>87.94</b>
	shuttle-2vs5	96.69	99.46	96.69	99.43	<b>99.59</b>	98.96
	vehicle2	95.62	95.69	96.32	94.80	88.95	<b>96.93</b>
	vowel0	95.23	95.65	96.23	95.59	86.74	<b>97.31</b>
	yeast1	70.60	70.58	70.20	69.09	70.15	<b>71.18</b>
	yeast5	96.43	94.98	96.46	94.79	<b>96.49</b>	95.49

$K$  is not the higher the better. The result in Fig. 2 reveals that the value of GM tends to decrease when the  $K$  is taken too large, which may be affected by noise samples. Thus, we recommend  $K \in [5, 12]$  as the default in applications. Moreover, we can also observe that the value of  $\gamma$  should not taken too small, since RIC may fail to fully mine the uncertain information in such a case. Thus,  $\gamma \in [0.02, 0.15]$  can be recommended in applications.

**Fig. 2.** Classification results of RIC with different parameters in the synthetic dataset.

### 3.4 Execution Time

The execution time in seconds of RIC and other comparison approaches on different datasets with SVM as shown in Table 4. We can see that the execution time of RIC is higher than other approaches since it needs to calculate a large number of distances between samples to obtain neighbors. In applications, the proposed RIC approach is more suitable for cases where high accuracy is required, whereas efficient computation is not a vital requirement.

**Table 4.** Execution time of different methods (In seconds)

Datasets	ROS	RUS	SMOTE	CBU	RUSBoost	RIC
ecoli2	0.1946	0.0512	0.1412	0.1800	2.9829	0.7967
ecoli4	0.1773	0.0589	0.1374	0.1389	2.7882	2.0548
led7digit	0.1885	0.0633	0.1487	0.1619	2.9539	1.5250
glass1	0.2001	0.0543	0.1380	0.1602	2.9692	0.4475
page-blocks0	216.1961	18.9831	186.3028	14.4645	3.0455	251.9016
shuttle-2vs5	71.1344	0.0591	70.4835	1.4666	2.2398	12.8517
vehicle2	14.9956	3.4566	14.5732	1.7628	3.0990	22.3558
vowel0	0.3269	0.0724	0.3118	0.2386	3.0903	1.7887
yeast1	0.2646	0.0633	0.2313	0.2744	3.0286	1.9574
yeast5	0.2919	0.0557	0.1942	0.2191	3.2067	4.5351

## 4 Conclusion

This paper proposes a reliability-based imbalanced data classification approach (RIC) with Dempster-Shafer theory. RIC considers not only the global optimization of different classifiers, but also the local optimization. Thus, we can obtain a more robust and reasonable performance for each test sample. The experiments on synthetic and several real imbalanced datasets have verified the effectiveness of RIC with respect to typical approaches. Moreover, we also investigate the influence of parameters on the classification performance of RIC. In the future, we will extend the application scope of RIC to other real-world tasks, such as network intrusion detection.

**Acknowledgments.** This work was supported by the National Natural Science Foundation of China under Grant U20B2067, Grant 61790552, and Grant 61790554; the Aeronautical Science Foundation of China under Grant 201920007001.

## References

1. Zhu, Z., Wang, Z., Li, D., Du, W.: Globalized multiple balanced subsets with collaborative learning for imbalanced data. *IEEE Trans. Cybern.* **52**(4), 2407–2417 (2022)
2. Chaovalitwongse, W.A., Fan, Y.J., Sachdeo, R.C.: On the time series  $k$ -nearest neighbor classification of abnormal brain activity. *IEEE Trans. Syst. Man Cybern. A Syst. Hum.* **37**(6), 1005–1016 (2007)
3. Kashef, R.: A boosted SVM classifier trained by incremental learning and decremental unlearning approach. *Expert Syst. Appl.* **167**, 114154 (2021)
4. He, H., Garcia, E.A.: Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.* **21**(9), 1263–1284 (2009)
5. Wang, N., Liang, R., Zhao, X., Gao, Y.: Cost-sensitive hypergraph learning with  $f$ -measure optimization, *IEEE Trans. Cybern.* (2021)
6. Seiffert, C., Khoshgoftaar, T.M., van Hulse, J., et al.: RUSBoost: a hybrid approach to alleviating class imbalance. *IEEE Trans. Syst. Man Cybern. A Syst. Hum.* **40**(1), 185–197 (2009)
7. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. *Ann. Statist.* **83**, 325–339 (1967)
8. Shafer, G.: *A Mathematical Theory of Evidence*. Princeton Univ. Press (1976)
9. Liu, Z., Pan, Q., Dezert, J., Han, J.W., He, Y.: Classifier fusion with contextual reliability evaluation. *IEEE Trans. Cybern.* **48**(5), 1605–1618 (2018)
10. Zhang, Z.W., Tian, H.P., Yan, L.Z., Martin, A., Zhou, K.: Learning a credal classifier with optimized and adaptive multiestimation for missing data imputation, *IEEE Trans. Syst. Man Cybern.* (2021)
11. Niu, J., Liu, Z., Lu, Y., Wen, Z.: Evidential combination of classifiers for imbalanced data. *IEEE Trans. Syst. Man Cybern.* (2022)
12. Zhang, Z.W., Ye, S.T., Zhang, Y.R., Ding, W.P., Wang, H.: Belief combination of classifiers for incomplete data. *IEEE-CAA J. Automatica Sin.* **9**(4), 652–667 (2022)
13. Grina, Fares, Elouedi, Zied, Lefevre, Eric: A preprocessing approach for class-imbalanced data using SMOTE and belief function theory. In: Analide, Cesar, Novais, Paulo, Camacho, David, Yin, Hujun (eds.) *IDEAL 2020. LNCS*, vol. 12490, pp. 3–11. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-62365-4\\_1](https://doi.org/10.1007/978-3-030-62365-4_1)
14. Niu, Jiawei, Liu, Zhunga: Imbalance data classification based on belief function theory. In: Denœux, Thierry, Lefèvre, Eric, Liu, Zhunga, Pichon, Frédéric. (eds.) *BELIEF 2021. LNCS (LNAI)*, vol. 12915, pp. 96–104. Springer, Cham (2021). [https://doi.org/10.1007/978-3-030-88601-1\\_10](https://doi.org/10.1007/978-3-030-88601-1_10)
15. Smets, P.: Constructing the pignistic probability function in a context of uncertainty. *Uncertainty Artif. Intell.* **5**, 29–39 (1990)
16. Zhang, Y., Liu, G., Luan, W.: An approach to class imbalance problem based on stacking and inverse random under sampling methods, In: 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), pp. 1–6. IEEE (2018)
17. Chawla, N.V., Bowyer, K.W., Hall, L.O., et al.: SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002)
18. Lin, W.C., Tsai, C.F., Hu, Y.H., et al.: Clustering-based undersampling in class-imbalanced data. *Inf. Sci.* **409**, 17–26 (2017)

# Author Queries

## Chapter 8

---

Query Refs.	Details Required	Author's response
AQ1	This is to inform you that corresponding author has been identified as per the information available in the Copyright form.	