British Journal of Mathematics & Computer Science
21(5): 1-9, 2017; Article no.BJMCS.32860
ISSN: 2231-0851

SCIENCEDOMAIN *international*
www.sciencedomain.org

Experimental Study on Class Imbalance Problem Using an Oil Spill Training Data Set

Xi Qin Ouyang[1*], Yuan Ping Chen[2] and Bing Hui Wei[3]

[1]Department of Mathematics, Ganzhou Teachers College, Ganzhou 341000, Jiangxi, China.
[2]Department of Foundation, Jiangxi Environmental Engineering Vocational College, Ganzhou 341000, Jiangxi, China.
[3]Department of Computer Science, Jiangxi University of Science and Technology, Ganzhou 341000, Jiangxi, China.

Authors' contributions

This work was carried out in collaboration between all authors. Author XQO designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Author YPC managed the literature searches. Author BHW managed the analyses of the study and revised the whole manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJMCS/2017/32860
Editor(s):
(1) Dijana Mosic, Department of Mathematics, University of Nis, Serbia.
Reviewers:
(1) Chunyang Gao, Guilin University of Electronic Technology, China.
(2) Merv Fingas, Spill Science Edmonton, Alberta T6W 1J6, Canada.
Complete Peer review History: http://www.sciencedomain.org/review-history/18614

Received: 20th March 2017
Accepted: 6th April 2017
Published: 13th April 2017

Short Research Article

Abstract

There is a paucity of research on one of the key issues in oil spill detection: the imbalanced training set learning problem. This paper performs experiments to show the influence of the imbalanced learning problem (ILP) on oil spill detection and devises a novel framework to tackle this problem. Experimental results show that an imbalanced training set degenerate the performance of oil spill detection, and our proposed framework achieves a better performance based on F-measure.

Keywords: Imbalanced problem; oil spill detection; data set.

**Corresponding author: E-mail: xixihappy123@163.com;*

ABBREVIATIONS

P: Positives; N: Negatives; NN: Neural Network; NB: Naive Bayes; SOM: Self-Organizing Map; MTD: Mega-Trend Diffusion; ILP: Imbalanced Learning Problem; SAR: Synthetic Aperture Radar; SMOTE: Synthetic Minority Oversampling Technique; MWMOTE: Majority Weighted Minority Oversampling Technique; MD: Mahalanobis Distance; CMD: Cost-sensitive Mahalanobis Distance; TSd: Standard Deviation of Target; BSd: Standard deviation of the background; TPR: True Positive Rate; TNR: True Negative Rate.

1 Introduction

Pollution resulting from oil spills in open sea and coastal waters is a major threat to ocean ecosystems. Detection and continuous monitoring of oil spills are important components of law enforcement efforts to minimize the impact that oil polluting events have on the ecosystem. Previous studies have shown that synthetic aperture radar (SAR) is effective in the detection and classification of oil spills. Oil spills appear as dark spots in SAR images. However, similar dark spots ("look-alikes"), resulting in misidentification, may arise from a range of unrelated meteorological and oceanographic phenomena. How to distinguish oil spills from "look-alikes" is definitely a hot topic.

Kubat and Brekke identified four key issues during their development of a machine learning component for an oil spill detection system [1,2]. One of the key issues is the imbalanced training set in which there are a great many more negative examples ("look-alikes") than positive examples (oil spills). Oil spill detection is an application where the classifier detects a rare, but important, event: "look-alikes" appear much more frequently. Imbalanced training set learning, namely designated imbalanced learning problem (ILP), moves the decision threshold, or classification boundary, toward the majority class, thereby causing a decrease in the generalization ability of machine learning algorithms. ILP is one of the key issues in oil spill detection; however, to our knowledge, how it affects detection performance and how to tackle it have not been reported in the literature.

This paper focuses on ILP and compares it with existing methods that deal with this problem using an oil spill training set. The contributions are two-folds: (1) to confirm that detection performance is affected by an imbalanced oil spill training data set and (2) to devise a novel, simple classification method. To descript clearly, we use an oil spill as a positive class, a minority class and a "look-alike" as a negative class, a majority class.

2 Methodology

2.1 State of art

To deal with ILP, many researchers have studied it mainly from three perspectives: sampling, cost-sensitive methods, and algorithm-based methods [3].

2.1.1 Sampling

Sampling methods include many different forms of generating and re-sampling. The main concept of the re-sampling methods is to balance class distribution by under-sampling the majority class examples, or over-sampling the minority class examples, or both. Under-sampling removes data from the majority of the class examples of the original data set and uses only some of the examples as the training set. Under-sampling, although readily providing a simple method for adjusting the balance of training set, may lead to an important loss of information. Kim proposed an under-sampling method based on a Self-Organizing Map (SOM), which is used in this work for comparison [4]. Over-sampling replicates selected examples of the minority class based on some strategies. Synthetic minority oversampling technique (SMOTE) [5] and

Mega-Trend Diffusion (MTD) function [6] are intelligent over-sampling methods that generate new examples for the minority class. But they ignored the information of majority. After then, majority weighted minority oversampling technique (MWMOTE) [7] was proposed to generate synthetic samples by combining minority and majority information. They create new examples artificially by interpolating the pre-existing minority examples. However, sampling-based method changes the true distribution of samples.

2.1.2 Cost-sensitive methods

Cost-sensitive methods are based on the concept of a cost matrix, which is considered a numerical representation of the penalty for misclassification [3]. In this work, the cost of misclassifying an oil spill as a "look-alike" is far greater than that of misclassifying a "look-alike" as an oil spill. Consequently, we attach a higher penalty to misclassifying an oil spill. This technique, in order to increase the probability of extracting an oil spill correctly, moves the decision threshold, or classifier boundary, toward the majority class. Some other methods introduce misclassification costs into the weight updating strategy used in AdaBoost [8].

2.1.3 Algorithm-based methods

The methods, at the algorithm level, include the following: adjusting the classification regularization [9]; changing the kernel function or corresponding weights [10]; considering only one class of information by ignoring the other classes of information [11,12]; hybrid framework with re-sampling and cost-sensitive methods; etc. Zong et al. proposed a weighted extreme learning machine method for imbalance learning [13].

2.2 Cost-sensitive Mahalanobis distance classification

Mahalanobis distance (MD), introduced by P. C. Mahalanobis in 1936, is a measure of the distance between a point, P, and a distribution, D. It is a multi-dimensional generalization of the idea of measuring the number of standard deviations P is from the mean of D [14]. Based on a given distribution, D, the MD of each point is computed as the input to the procedure that follows. The classifier performs based on the MD of each data point. Both Fiscella and Zhou used this classification in oil spill detection, while the difference of our method in this work is at extracting normal group and defining the threshold [15,16].

Cost-sensitive Mahalanobis distance (CMD) classification combines a cost-sensitive method, which takes into account a different penalty for misclassifying different classes. The additional process defines a threshold that is suitable as a classifier. The framework is introduced in detail in the following sections.

2.2.1 Extracting normal group

Based on a training data set that includes oil spills and "look-alike" examples, we compute the Euclidean distance between any two examples. We take the five (or other threshold) nearest neighbor examples as the neighborhood for each "look-alike" example. If there is no "look-alike" example in the neighborhood, we consider the corresponding "look-alike" example as noisy. Consequently, the noisy "look-alike" examples are removed, and the remaining "look-alike" examples are taken as a normal group. Note that how to define the neighborhood is based on the application in different cases. Here, it is defined as five.

This technique has a greater probability to separate an oil spill from a "look-alike" and defines a distribution that is more representative of "look-alike" examples. Noisy data affects the distribution; thus, it must be removed from the normal group.

2.2.2 Computing Mahalanobis distance

Here, we take "look-alikes" as a normal group, which are derived from a distribution space; however, oil spills are abnormal examples, which are far removed from the normal group.

Assume that $\mathbf{X} \in \mathrm{R}^{m*n}$, where m and n denote the dimension and number of features respectively, is an observation set that is derived from the normal group. $\boldsymbol{\mu} \in \mathrm{R}^{m}$ is the mean vector and $\mathbf{S} \in \mathrm{R}^{m \times m}$ is the covariance matrix of X. The MD of a point, $\boldsymbol{y} \in \mathrm{R}^{m}$, from the normal group is represented as follows:

$$\mathrm{D}(\boldsymbol{y}, \mathbf{X}) = \sqrt{(\boldsymbol{y} - \boldsymbol{\mu})^{\mathrm{T}} \mathbf{S}^{-1} (\boldsymbol{y} - \boldsymbol{\mu})} \tag{1}$$

Based on this representation, all data (including oil spills and “look-alikes”) are changed into another measurement scale. Fig. 1 shows the results derived from a trial. The map of an oil spill (red ‘+’ in Fig. 1) is far removed from that of a “look-alike” group (blue ‘o’ in Fig. 1).

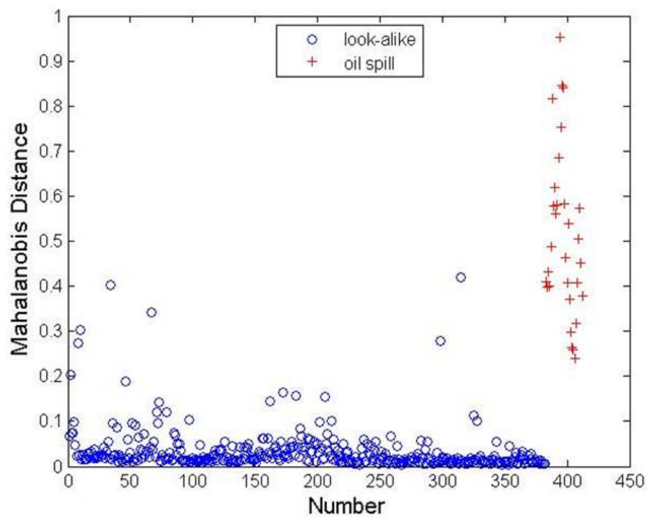


Fig. 1. Mahalanobis distance of examples including oil spill and “look-alike” based on extracted normal group

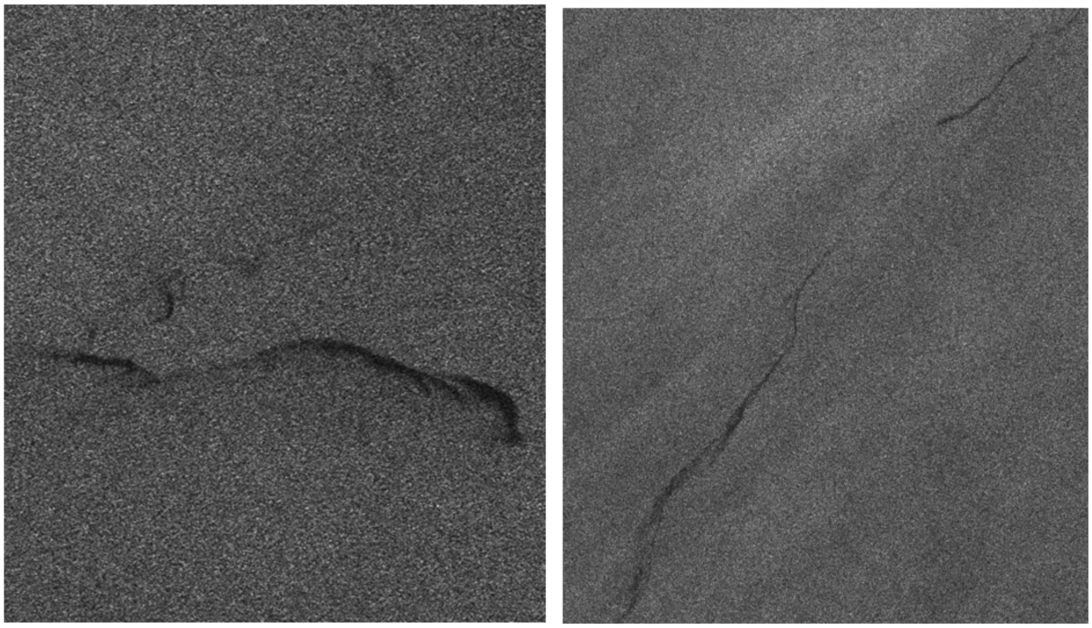
2.2.3 Defining decision threshold

We classify test data, based on its MD, by comparing the data to a given threshold. Here, we introduce cost-sensitivity into defining the threshold. The cost of misclassifying an example is defined based on its significance. There are many cost versions in the literature. We define the cost of a class example as one divided by the number of class examples, which shows that the total significance of any given class is equal to one unit. Then, we search the cost from the mean MD of “look-alikes” to the mean MD of oil spill training data. The minimum cost corresponds to the threshold we want to find, i.e. the best boundary to segment the two classes.

3 Experiments

3.1 Data set

The dataset used in this study is derived from RADARSAT-1 ScanSAR narrow beam images with a swath of 300 km and a spatial resolution of 50 m, and covers vast Pacific and Atlantic coastal areas [17]. The dataset used comprises fourteen features of 412 oil spills and “look-alikes”. There are thirty oil spills and 382 “look-alikes.” Fig. 2 shows examples of oil spill and “look-alike”. By visually discerning the gray tone difference between the dark-spots and the background, we delineated dark-spot boundaries; therefore, we need not introduce dark-spot detection.



(a) **(b)**

Fig. 2. The dark spots are the examples of (a) "look-alike" (acquired on Sep. 17, 2013, located at 50°43′ N, 53°25′ W), and (b) oil spill (acquired on Jul. 23, 2009, located at 42°50′ N, 56°38′ W).

Given the dark-spots in pixel-format, features must be extracted as input for the classifiers. For each of the dark spots, the following features, proposed in this paper, are computed:

3.1.1 Physical and textural features

(1) Target area A, (2) Target perimeter P, (3) Complexity measure, C=P2/A, and (4) Spreading measure S, i.e. the ratio between target width and length.

3.1.2 Geometric features

(1) Standard deviation of target (TSd), (2) Mean intensity of the background area (BM), and (3) Standard deviation of the background (BSd).

3.1.3 Contrast with background

(1) Power to mean ratio contrast (Tpm/Bpm) where Bpm=BSd/BM, Tpm=TSd/TM and TM represent the mean intensity of the target, (2) ConRaSd =Sd/BSd, (3) ConLa, defined as the ratio between the TM and the mean intensity value of a window centered at the region, (4) ConMax, defined as the difference between the BM and the lowest value inside the target, (5) ConSm=(Nt/Gt)/(Nb/Gb) where Nt is the number of target pixels, Gt is the sum of the gradient values of target pixels, Nb is the number of background window pixels, Gb is the sum of the background window gradient values, (6) Maximum gradient value of the dark-spot border area, GMax. The gradients are computed by the Sobel operator, (7) Standard deviation of the border gradient values, GSd. More details about the data and features of dark-spot refer to Xu et al. [17].

3.2 Evaluation metric

To evaluate classifiers on the imbalanced data sets, we use a true positive rate (TPR), a true negative rate (TNR), and F metrics. Table 1 shows the corresponding confusion matrix, where Positive denotes oil spill and Negative denotes "look-alike". TPR, TNR and F metric (derived from the version in Sokolova et al.) [18] are defined as follows:

$$\mathrm{TPR}=\frac{\mathrm{TP}}{\mathrm{TP+FP}} \qquad \mathrm{TNR}=\frac{\mathrm{TN}}{\mathrm{TN+FN}} \qquad \mathrm{F}=\frac{2\times Precision\times Recall}{Precision+Recall}=\frac{2\times\frac{\mathrm{TP}}{\mathrm{TP+FP}}\times\frac{\mathrm{TP}}{\mathrm{TP+FN}}}{\frac{\mathrm{TP}}{\mathrm{TP+FP}}+\frac{\mathrm{TP}}{\mathrm{TP+FN}}} \tag{2}$$

where Precision, Recall are two popular evaluation metrics, and TP, TN, FP, FN are shown in Table 1.

Table 1. Confusion matrix for performance evaluation

P (oil spills = Positives)	N (look-alikes = Negatives)
TP (True Positives)	FN (False Negatives)
FP (False Positives)	TN (True Negatives)

Intuitively, TPR and TNR are measure of exactness for actually labeled correctly. The F measure metric combines exactness and completeness as a measure of classification effectiveness. We use the TPR and TNR metrics to illustrate intuitively each method in dealing with an imbalanced training set. We use the F measure to evaluate the performance of each classification method.

3.3 Experiments results and discussion

Our experiments include the following three parts. In the first two parts, we carry out the following classifications: (1) k-Nearest Neighbor (10-NN); (2) decision tree (C4.5); (3) Naive Bayes (NB); (4) Neural Network (NN) with the code in Matlab7.14 where parameters are set optimally; (5) SVM with the library, Libsvm, written by Chih-Chung Chang and Chih-Jen Lin; (6) AdaBoost introduced in Pang et al. [9]. We employ a three-fold cross-validation technique for performance estimation. In each trial, we divide both the oil spill dataset and the "look-alike" dataset into three subsets of equal size, select two subsets of oil spill and two subsets of "look-alike" as training set and the remainder as testing set.

3.3.1 First part, combining re-sampling methods into classifier

The first part is to carry out an under-sampling method and two over-sampling methods based on the following six classifications that are the population in machine learning: k-NN, SVM, NB, NN, C4.5 and AdaBoost methods. We introduce three sampling methods as follows, that is SOM, MTD and MWMOTE.

SOM is an unsupervised algorithm. First, it clusters the original majority class examples. Then, to remove the examples which close to the centroid of each cluster by given probability. Consequently, the remainder majority class examples will be taken as the most representative examples and as the input to classifier. This method reduces the number of majority class examples, but, it can't change the distribution of boundary examples in majority class. We use it to extract, from the training set, the samples that best represent the majority of the class samples.

Li et al. [19] proposed the MTD function to deal with the small data set problem for scheduling strategies in early flexible manufacturing systems. The main purpose of the MTD function is to reduce the number of data in the majority class and generate synthetic examples from the minority class to solve the ILP. The MTD function first extracts only the part of the majority examples that is most representative of its category. Then, the MTD function generates synthetic examples based on a linear combination randomly selected from among the minority class examples. Consequently, the number of examples between the majority and minority classes is transferred to a more balanced level.

MWMOTE first identifies the hard-to-learn informative minority class examples and assigns them weights according to their Euclidean distance from the nearest majority class samples. Then, using a clustering approach, it generates synthetic samples from the weighted informative minority class examples. In such a way, all the generated examples lie inside some minority class cluster.

We carry out the classifiers aforementioned and Table 2 shows the results. Numbers 1-6 in Table 2 show the results of six classifiers on the observed original training data. The TPR values are the same for all six, but the TNR values are different. For the F measure, C4.5 and AdaBoost classifications exhibit the best results. SVM shows a somewhat poorer performance. Numbers 7-24 show the results based on the three sampling methods. For SOM-based under-sampling, the performance of C4.5 and AdaBoost are invariant with that based on the original training data. Due to a reduced majority class, the other four classifiers show worse performance; especially for k-NN, performance declines severely. Therefore, this under-sampling method does not affect the C4.5 and AdaBoost classifiers, and it is unsuitable for the other four classifiers.

Table 2. Performance of classification

No	Methods	TPR	TNR	F	No	Methods	TPR	TNR	F
1	KNN	0.77	0.82	0.43	18	AdaBoost + MTD	0.87	0.46	0.30
2	SVM	0.77	0.94	0.62	19	KNN + MWMOTE	0.77	0.81	0.43
3	NB	0.77	0.79	0.41	20	SVM + MWMOTE	0.77	0.92	0.59
4	NN	0.77	0.80	0.42	21	NB + MWMOTE	0.77	0.74	0.38
5	C4.5	0.77	0.95	0.64	22	NN + MWMOTE	0.77	0.77	0.40
6	AdaBoost	0.77	0.95	0.64	23	C4.5 + MWMOTE	0.77	0.93	0.61
7	KNN + SOM	0.40	0.66	0.18	24	AdaBoost + MWMOTE	0.77	0.95	0.64
8	SVM + SOM	0.73	0.72	0.35	25	KNN + cost-sensitive	0.83	0.80	0.44
9	NB + SOM	0.77	0.79	0.41	26	SVM + cost-sensitive	0.87	0.92	0.61
10	NN + SOM	0.77	0.73	0.37	27	NB + cost-sensitive	0.83	0.76	0.42
11	C4.5 + SOM	0.77	0.95	0.64	28	NN + cost-sensitive	0.83	0.78	0.43
12	AdaBoost + SOM	0.77	0.95	0.64	29	C4.5 + cost-sensitive	0.87	0.92	0.61
13	KNN + MTD	0.90	0.47	0.32	30	AdaBoost + cost-sensitive	0.87	0.92	0.61
14	SVM + MTD	0.97	0.19	0.28	31	One class SVM(negative)	0.40	0.95	0.40
15	NB + MTD	0.90	0.49	0.33	32	One class SVM(positive)	0.97	0.19	0.28
16	NN + MTD	0.93	0.41	0.32	33	Extreme machine learning	0.77	0.75	0.38
17	C4.5 + MTD	0.87	0.46	0.30	34	Cost-sensitive MD	0.97	0.91	0.67

For MTD-based over-sampling, as expected, the TPR of all classifiers increase; however, their TNRs decrease dramatically. Consequently, the F measures of these classifiers decrease in different degrees. By removing some majority examples and generating some synthetic examples based on minority examples, this over-sampling method makes the decision threshold move toward a majority; thus, this over-sampling method increases the TPR, while adding some noisy data to the training set.

WMWOTE tackles this problem, and generates synthetic examples based on only the minority examples that pre-removed noisy data. However, the TPRs of WMWOTE-based classifiers are invariant, and their TNRs decrease. The degree of the decrease in their TNRs is less than that of the MTD-based classifiers.

These experimental results show that imbalanced oil spill learning is affected, more or less. To increase the probability of correctly classifying an oil spill, re-sampling methods move the decision threshold toward a majority class. But, at the same time, the procedure decreases the probability of classifying “look-alikes.” If re-sampling processing reduces the total cost of mis-classification, the processing is feasible.

3.3.2 Second part, combining cost-sensitive into classifier

Secondly, to show the influence of cost-sensitive processing for ILP on an oil spill training set, these classifiers are implemented with a cost-sensitive method. We add a cost matrix to the classifier’s initial input. The same as previously done in Section 3, the costs of an oil spill and a “look-alike” are set at one divided by the number of oil spill examples and one divided by the number of "look-alike" examples, respectively. In other words, we assume that the total importance of one class is equal to that of another. It seems reasonable that, the fewer the number of examples in a class, the greater the importance of each example.

For the cost-sensitive classifiers, experimental results show the similar case in MTD-Based methods that the increase of TPR is at the expense of its TNR. Also, the F measures of these classifiers varied in different degrees. This case illustrates that cost-sensitivity must be taken into account in imbalanced oil spill learning. Numbers 25-30 in Table 2 show the best performance for the cost of an oil spill, in the range 1 to 12.7, which maximum upper bound is an imbalanced ratio equal to the number of "look-alike" examples divided by the number of oil spill examples.

3.3.3 Third part, other algorithms for ILP

Thirdly, the following other algorithms are carried out for ILP: one-class SVM, extreme machine learning, and our proposed framework. The significance of performance differences among classifiers is statistically tested. One-class SVM takes only one class of information (oil spill or "look-alike") as classifier input, while ignoring other classes of information. We performed this classifier with two strategies that use as input oil spill and "look-alike," respectively. The results shown in Numbers 31-32 in Table 2, demonstrate that the correct ratio of the class that is taken as classifier input is much greater than that of another class that is ignored in the training procedure. The F measure of the two strategies is low. Extreme machine learning performance shown in Number 33 is poor.

Compared with all the above mentioned classifiers including Sampling, Cost-sensitive, algrithm-based methods, our proposed framework, whose results showed in Number 34 in Table 2, achieves the best performance based on F-measure. On the oil spill data set, the performance of Sampling and Cost-sensitive classification is not superior to that based on original classification. The experiment on Cost-sensitive MD showed that most oil spills are detected correctly, and few part of "look-alike" examples are classified as oil spills, mainly because the higher cost of an oil spill pushes the decision threshold toward the majority class.

4 Conclusion

In this paper, we performed many classifications to study oil spill detection using imbalanced learning methods. We analyzed the influence an imbalanced training set has in detecting oil spills. We drew the conclusion that we must consider oil spill importance as an additional process in classification. Furthermore, regarding imbalanced oil spill learning, we devised a novel framework that achieves its best performance based on and F measures. However, due to the limitation of the oil spill training data, we didn't conduct more experiments to further identify the advantages of our proposed framework.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Kubat M, Holte RC, Matwin S. Machine learning for the detection of oil spills in satellite radar images. Machine Learning. 1998;30(2-3):195-215.

[2] Brekke C, Solberg AH. Oil spill detection by satellite remote sensing. Remote Sensing of Environment. 2005;95(1):1-13.

[3] He HB, Garcia EA. Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering. 2009;21(9):1263-1284.

[4] Kim MS. An effective under-sampling method for class imbalance data problem. Isis Proceedings of Symposium on Advanced Intelligent Systems. 2007:825-829.

[5] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research. 2002;16:321-357.

[6] Li DC, Liu CW, Hu SC. A learning method for the class imbalance problem with medical data sets. Computers in Biology and Medicine. 2010;40(5):509-518.

[7] Barua S, Islam MM, Yao X, Murase K. MWMOTE-majority weighted minority oversampling technique for imbalanced data set learning. IEEE Transactions on Knowledge and Data Engineering. 2014;26(2):405-425.

[8] Sun YM, Kamel MS, Wong AK, Wang Y. Cost-sensitive boosting for classification of imbalanced data. Pattern Recognition. 2007;40(12):3358-3378.

[9] Pang SN, Zhu L, Chen G, Sarrafzadeh A, Ban T, Inoue D. Dynamic class imbalance learning for incremental LPSVM. Neural Networks. 2013;44:87-100.

[10] Maratea A, Petrosino A, Manzo M. Adjusted F-measure and kernel scaling for imbalanced data learning. Information Sciences. 2014;257:331-341.

[11] Manevitz LM, Yousef M. One-class SVMs for document classification. Journal of Machine Learning Research. 2001;2:139-154.

[12] Zhuang L, Dai HH. Parameter estimation of one-class SVM on imbalance text classification. Paper presented at: Conference of the Canadian Society for Computational Studies of Intelligence; 2006.

[13] Zong WW, Huang GB, Chen YQ. Weighted extreme learning machine for imbalance learning. Neurocomputing. 2013;101:229-242.

[14] Mahalanobis PC. On the generalized distance in statistics. Proceedings of the National Institute of Sciences (Calcutta). 1936;2:49-55.

[15] Fiscella B, Giancaspro A, Nirchio F, Pavese P, Trivero P. Oil spill detection using marine SAR images. International Journal of Remote Sensing. 2000;21(18):3561-3566.

[16] Zhou H, Chen P. Oil Spill detection using mahalanobis distance based on synthetic aperture radar images. Computer Engineering and Applications. 2006;47(18):195-197 (In Chinese).

[17] Xu LL, Li J, Brenning A. A comparative study of different classification techniques for marine oil spill identification using RADARSAT-1 imagery. Remote Sensing of Environment. 2014;141:14-23.

[18] Sokolova M, Japkowicz N, Szpakowicz S. Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation. Paper presented at: Australasian Joint Conference on Artificial Intelligence; 2006.

[19] Li DC, Wu CS, Tsai TI, Lina YS. Using mega-trend-diffusion and artificial samples in small data set learning for early flexible manufacturing system scheduling knowledge. Computers & Operations Research. 2007;34(4):966-982.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/18614