Integrating Selective Pre-processing of Imbalanced Data with Ivotes Ensemble

Jerzy Błaszczyński, Magdalena Deckert, Jerzy Stefanowski, Szymon Wilk

Institute of Computing Science, Poznań University of Technology,
60-965 Poznań, Poland
{jerzy.blaszczynski, magdalena.deckert, jerzy.stefanowski, szymon.wilk}@cs.put.poznan.pl

Abstract. In the paper we present a new framework for improving classifiers learned from imbalanced data. This framework integrates the SPIDER method for selective data pre-processing with the Ivotes ensemble. The goal of such integration is to obtain improved balance between the sensitivity and specificity for the minority class in comparison to a single classifier combined with SPIDER, and to keep overall accuracy on a similar level. The IIvotes framework was evaluated in a series of experiments, in which we tested its performance with two types of component classifiers (tree- and rule-based). The results show that IIvotes improves evaluation measures. They demonstrated advantages of the abstaining mechanism (i.e., refraining from predictions by component classifiers) in IIvotes rule ensembles.

1 Introduction

Learning classifiers from imbalanced data has received a growing research interest in the last decade. In such data, one of the classes (further called a *minority class*) contains significantly smaller number of objects than the remaining *majority classes*. The imbalanced class distribution causes difficulties for the majority of learning algorithms because they are biased toward the majority classes and objects from the minority class are frequently misclassified, what is not acceptable in many practical applications.

Several methods have been proposed to deal with learning from imbalanced data (see [5, 6] for reviews). These methods can be categorized in two groups. The first group includes classifier-independent methods that rely on transforming the original data to change the distribution of classes, e.g., by re-sampling. The other group involves modifications of either learning or classification strategies.

In this paper, we focus on re-sampling techniques. The two well known methods are SMOTE for selective over-sampling of the minority class [3], and NCR for removing objects from the majority classes [9]. Stefanowski and Wilk also proposed a new method to selective pre-processing combining filtering and over-sampling of imbalanced data (called SPIDER) [12]. Experiments showed that it was competitive to SMOTE and NCR [13]. Unfortunately, for some data sets the improvement of the sensitivity for the minority class was associated with too large decrease of specificity for this class (it translated into worse recognition of objects from the majority classes). It affects SPIDER and other methods included in the experiment.

In our opinion it is an undesirable property as in many problems it is equally important to improve sensitivity of a classifier induced from imbalanced data and to keep its specificity and overall accuracy at an acceptable level (i.e., both measures should not deteriorate too much comparing to a classifier induced from data without pre-processing). We claim that in general there is a kind of trade off between these measures and too large drop of specificity or accuracy may not be accepted. Thus, our goal is to modify SPIDER in a way that would improve this trade-off.

To achieve it we direct out attention to *adaptive ensemble classifiers* which iteratively construct a set of component classifiers. Such classifiers optimize the overall accuracy, by iteratively learning objects which were difficult to classify in previous iterations. However, as these objects are sampled from the original learning set which is predominated by the majority classes, even misclassified objects may be still biased toward these classes. Our proposition to overcome this problem is using the SPIDER method to transform each sample in succeeding iterations. It should increase the importance of the minority class objects in learning each component classifier. As an ensemble we decided to consider the Ivotes approach introduced by Breiman in [2], as it is already based on a kind of focused sampling of learning objects. Moreover, we have already successfully applied this ensemble with the MODLEM rule induction algorithm [10, 11] and we think its classification strategy could be biased toward the minority class with so-called abstaining [1].

A similar idea of using adaptive ensembles was followed in the SMOTEBoost algorithm [4], where the basic SMOTE method was successfully integrated with changing weights of objects inside the AdaBoost procedure. Results reported in the related literature show that Ivotes gives similar classification results as boosting, therefore we hope that our solution will also work efficiently.

The main aim of this paper is to present the new framework for dealing with imbalanced data based on incorporating SPIDER into the Ivotes ensemble. We evaluate its performance experimentally on several imbalanced data sets and we compare it to the performance of single classifiers combined with SPIDER. We consider tree-based and rule-based classifiers induced by the C4.5 and the MODLEM algorithms respectively, as according to previous studies they are sensitive to the class imbalance [12, 13].

2 Related Works

In this section we concentrate on these re-sampling methods that are most related to our study – for reviews of other approaches see [5, 6]. Kubat and Matwin in their paper on one-side sampling claimed that characteristics of mutual positions of objects is a source of difficulty [8]. They focus attention on *noisy* objects located inside the minority class and *borderline* objects. Such objects from the majority classes are removed while keeping the minority class unchanged. Another approach to focused removal of objects from the majority classes is the NCR method introduced in [9], which uses the Edited Nearest Neighbor Rule (ENNR) and removes these objects from the majority classes that are misclassified by its k nearest neighbors. The best representative of focused over-sampling is SMOTE that over-samples the minority class by creating new synthetic objects in the k-nearest neighborhood [3].

However, some properties of these methods are questionable. NCR or one-side-sampling may remove too many objects from the majority classes. As a result improved sensitivity is associated with deteriorated specificity. Random introduction of synthetic objects by SMOTE may be questionable or difficult to justify in some domains, where it is important to preserve a link between the original data and a constructed classifier. Moreover, SMOTE may blindly "over-generalize" the minority class area without checking positions of the nearest objects from the majority classes, thus increasing overlapping between classes.

Following this criticism Stefanowski and Wilk introduced SPIDER – a new method for selective pre-procesing [12]. It combines removing these objects from the majority classes that may result in misclassification of objects from the minority class, with local over-sampling of these objects from the minority class that are "overwhelmed" by surrounding objects from the majority classes. On the one hand, such filtering is less greedy than the one employed by NCR, and on the other hand, over-sampling is more focused that this used by SMOTE. SPIDER offers three filtering options that impact modification of the minority class and result in changes of increasing degree and scope: *weak amplification*, *weak amplification and relabeling*, and *strong amplification*. More detailed description is given in Section 3.

Finally, let us note that various re-sampling techniques were integrated with ensembles. The reader is referred to a review in [6] that besides SMOTEBoost describes such approaches as DataBoost-IM or special cost-sensitive modifications of AdaBoost.

3 Proposed Framework

Our framework combines selective pre-processing (SPIDER) with an adaptive ensemble of classifiers. Such ensembles are able to adapt to objects that are difficult to learn in succeeding iterations. Such difficult objects from the majority class could be especially important when learning from imbalanced data. We decided to use Ivotes [2] as the ensembles due to reasons given in Section 1. We propose to incorporate SPIDER inside this ensemble to obtain a classifier more focused on minority class. However, due to the construction of the ensemble and its general controlling criterion (accuracy) we still expect that it should sufficiently balance the sensitivity and specificity for the minority class.

The resulting Imbalanced Ivotes (shortly called IIvotes) algorithm is presented in Figure 1. In each iteration, IIvotes creates a new training set from LS by *importance sampling*. The rationale for the importance sampling is that the new training set will contain about equal numbers of incorrectly and correctly classified objects. In this sampling an object is randomly selected with all objects having the same probability of being selected. Then it is classified by an out-of-bag classifier (i.e., ensemble composed of all classifiers which were not learned on the object). If the object is misclassified then it is selected into the new training set S_i. Otherwise, it is sampled into S_i with probability $\frac{e(i)}{1-e(i)}$, where $e(i)$ is a generalization error. Sampling is repeated until n objects are selected. Each S_i is processed by SPIDER. In each iteration, $e(i)$ is estimated by out-of-bag classifier. IIvotes iterates until $e(i)$ stops decreasing.

The SPIDER method is presented in Figure 2. In the pseudo-code we use the following auxiliary functions (in all these functions we employ the heterogeneous value distance metric (HVDM) [9] to identify the nearest neighbors of a given object):

- `correct`(S, x, k) – classifies object x using its k-nearest neighbors in set S and returns true or false for correct and incorrect classification respectively.
- `flagged`(S, c, f) – identifies and returns a subset of objects from set S that belong to class c that are flagged as f.
- `knn`(S, x, k, c, f) – identifies and returns these objects among the k-nearest neighbors of x in set S that belong to class c and are flagged as f.
- `amplify`(S, x, k, c, f) – amplifies object x by creating its |`knn`(S, x, k, c, f)| copies and adding it to set S (where $|.|$ denotes the cardinality of a set).

SPIDER consists of two main phases – *identification* and *pre-processing*. In the first phase it identifies the "local" characteristics of objects following the the idea of ENNR [9], flags them appropriately, and marks questionable objects from c_{maj} for possible removal. In the second phase , depending on the pre-processing option SPIDER amplifies selected objects from c_{min}, relabels selected questionable objects from c_{maj} (i.e., their class is changed to c_{min}), and finally removes remaining questionable objects from c_{maj} from a resulting data set. Much more thorough description of the method is provided in [12, 13].

Let us remark that Ivotes ensembles proved to improve their performance in terms of predictive accuracy with component classifiers that are able to abstain (i. e., they do not classify objects when they are not sufficiently certain) [1]. We are interested in checking whether abstaining could also help in classifying objects from the minority class. According to our previous experience [1], abstaining can be implemented by changing classification strategies inside rule ensembles (by refraining from prediction, when the new object is not precisely covered by rules in the component classifiers).

4 Experiments

The main aim of our experiments was to evaluate the ability of the new IIvotes framework to balance the recognition of minority and majority classes. Thus, we compared the performance of IIvotes with three pre-processing options for SPIDER (weak, relabel and strong – see Figure 2) to the performance of single

Algorithm 1: IIvotes

Input : LS – learning set; TS – testing set; n – size of learning data set; LA – learning algorithm; c_{min} – the minority class; k – the number of nearest neighbors; *opt* – pre-processing option of SPIDER
Output: C^* final classifier

```
Learning phase
while e(i) < e(i − 1) do
    S_i := importance sample of size n from LS
    S_i := SPIDER (S_i, c_min, k, opt) {selective pre-processing of S_i}
    C_i := LA (S_i) {construct a base classifier}
    e(i) := estimate generalization error by out-of-bag classifier
    i := i + 1
Classification phase
foreach x ∈ TS do
    C*(x) = arg max_X Σ_{i=1}^{T} (C_i(x) = X) {the class with maximum number of
    votes is chosen as a final label for x}
```

Algorithm 2: SPIDER

Input : DS – data set; c_{min} – the minority class; k – the number of nearest neighbors; *opt* – pre-processing option (weak = weak amplification, relabel = weak amplification and relabeling, strong = strong amplification)
Output: pre-processed DS

```
c_maj := an artificial class combining all the majority classes in DS
Identification phase
foreach x ∈ DS do
    if correct(DS, x, k) then flag x as safe
    else flag x as noisy
RS := flagged(DS, c_maj, noisy)
Pre-processing phase
if opt = weak ∨ opt = relabel then
    foreach x ∈ flagged(DS, c_min, noisy) do amplify(DS, x, k, c_maj, safe)
    if opt = relabel then
        foreach x ∈ flagged(DS, c_min, noisy) do
            foreach y ∈ knn(DS, x, k, c_maj, noisy) do
                change classification of y to c_min
                RS := RS \ {y}
else // opt = strong
    foreach x ∈ flagged(DS, c_min, safe) do amplify(DS, x, k, c_maj, safe)
    foreach x ∈ flagged(DS, c_min, noisy) do
        if correct(DS, x, k + 2) then amplify(DS, x, k, c_maj, safe)
        else amplify(DS, x, k + 2, c_maj, safe)
DS := DS \ RS
```

classifiers combined with the same SPIDER pre-processing. Moreover, for comprehensive comparison we introduced the following baseline classifiers (further denoted as base) – Ivotes ensembles for IIvotes ensembles and single classifiers without any pre-processing for single classifiers with SPIDER.

We constructed all classifiers with two learning algorithms – C4.5 (J48 from WEKA) for decision trees and MODLEM for decision rules (MODLEM is described in [10, 11] and applied together with Grzymala's LERS strategy for classifying new objects [7]). Both algorithms were run without prunning to get more precise description of the minority class. SPIDER was used with $k = 3$ neighbors and the size of sample n in IIvotes was set to 50% based on our experience from previous experiments. In case of rule ensembles, besides the basic construction, we additionally tested a version with abstaining of component classifiers [1]. All algorithms were implemented in Java using WEKA.

Table 1. Characteristics of data sets

Data set	Objects	Attributes	Minority class	Imbalance ratio
abdominal-pain	723	13	positive	27.94%
balance-scale	625	4	B	7.84%
breast-cancer	286	9	recurrence_events	29.72%
bupa	345	6	sick	42.03%
car	1728	6	good	3.99%
cleveland	303	13	positive	11.55%
cmc	1473	9	long-term	22.61%
ecoli	336	7	imU	10.42%
german	666	20	bad	31.38%
haberman	306	3	died	26.47%
hepatitis	155	19	die	20.65%
pima	768	8	positive	34.90%
transfusion	748	4	yes	23.80%

The experiments were carried out on 13 data sets listed in Table 1. They either came from the UCI repository[1] or from our medical case studies (abdominal pain). We selected data sets that were characterized by varying degrees of imbalance and that were used in other related works.

All experiments were run with a stratified 10-fold cross-validation repeated five times. Besides recording average values of sensitivity, specificity and overall accuracy we also used G-mean – geometric mean of sensitivity and specificity – to evaluate the balance between these two measures. G-mean (GM in short) was proposed in [8] as a replacement for overall accuracy to maximize the recognition of the minority and majority classes, and since then it has been used in multiple studies on learning from imbalanced data. GM for tree- and rule-based classifiers

[1] http://www.ics.uci.edu/~mlearn/MLRepository.html

are presented in Table 2 and 3. Moreover, in Table 4 we show GM for IIvotes rule ensembles with abstaining.

Table 2. GM for tree-based classifiers

Data set	Single C4.5				Ivotes / IIvotes + C4.5			
	Base	Weak	Relabel	Strong	Base	Weak	Relabel	Strong
abdominal-pain	0.7812	0.7859	0.7807	0.7919	0.8052	0.8216	0.8239	0.8157
balance-scale	0.0249	0.2648	0.3646	0.2562	0.0881	0.4584	0.3827	0.5232
breast-cancer	0.5308	0.5487	0.5824	0.5602	0.5467	0.6068	0.5868	0.5683
bupa	0.6065	0.6032	0.5628	0.6037	0.6635	0.6804	0.7019	0.6612
car	0.8803	0.9261	0.8603	0.9111	0.8093	0.9149	0.8945	0.9171
cleveland	0.3431	0.4531	0.5052	0.4079	0.2759	0.4411	0.3914	0.4896
cmc	0.5533	0.6378	0.6175	0.6310	0.5813	0.6620	0.6439	0.6547
ecoli	0.6924	0.7728	0.7788	0.7852	0.7443	0.8383	0.8122	0.8462
german	0.5828	0.6114	0.6113	0.6086	0.5947	0.6738	0.6615	0.6662
haberman	0.5375	0.6089	0.6083	0.6118	0.4750	0.6256	0.6085	0.6167
hepatits	0.5386	0.5971	0.6518	0.5534	0.7115	0.7642	0.7466	0.7422
pima	0.6949	0.6978	0.7046	0.6986	0.7255	0.7401	0.7340	0.7343
transfusion	0.5992	0.6276	0.6317	0.6252	0.5181	0.6492	0.6523	0.6309

Table 3. G-means for rule-based classifiers (rule ensembles without abstaining)

Data set	Single MODLEM				Ivotes / IIvotes + MODLEM			
	Base	Weak	Relabel	Strong	Base	Weak	Relabel	Strong
abdominal-pain	0.7731	0.7968	0.7914	0.7946	0.7933	0.8321	0.8183	0.8278
balance-scale	0.0000	0.1913	0.1613	0.1722	0.0634	0.1125	0.0729	0.1454
breast-cancer	0.5008	0.5612	0.5104	0.5687	0.4748	0.5571	0.5462	0.5837
bupa	0.6502	0.5969	0.6725	0.5989	0.6703	0.6800	0.7002	0.6920
car	0.8978	0.9547	0.9404	0.9489	0.9021	0.9722	0.9638	0.9779
cleveland	0.3292	0.4360	0.3738	0.4673	0.1063	0.3307	0.2364	0.3628
cmc	0.5171	0.6320	0.5770	0.6218	0.5304	0.6660	0.6029	0.6575
ecoli	0.6502	0.7736	0.6655	0.7763	0.6140	0.7879	0.7233	0.7969
german	0.5499	0.6147	0.5719	0.6337	0.5133	0.6272	0.5838	0.6382
haberman	0.4588	0.5382	0.4790	0.5702	0.4345	0.5403	0.4807	0.5570
hepatits	0.6140	0.6861	0.6082	0.6482	0.6142	0.6637	0.6702	0.6817
pima	0.6576	0.7190	0.6832	0.7148	0.6510	0.7356	0.6944	0.7271
transfusion	0.5128	0.6153	0.5422	0.6103	0.4848	0.6100	0.5693	0.6239

Table 4. GM for rule ensembles with abstaining

Data set	Ivotes / IIvotes + MODLEM Base	Weak	Relabel	Strong
abdominal-pain	0.7995	0.8345	0.8284	0.8400
balance-scale	0.0625	0.1637	0.0878	0.2470
breast-cancer	0.5203	0.5776	0.5716	0.5886
bupa	0.7045	0.7058	0.7124	0.6933
car	0.9426	0.9743	0.9780	0.9834
cleveland	0.2361	0.4028	0.3232	0.4420
cmc	0.5630	0.6684	0.6353	0.6709
ecoli	0.7098	0.8077	0.7706	0.8245
german	0.6055	0.6852	0.6512	0.6885
haberman	0.4944	0.5704	0.5044	0.5625
hepatits	0.6759	0.7047	0.7005	0.7240
pima	0.7049	0.7507	0.7306	0.7430
transfusion	0.5331	0.6212	0.5851	0.6324

Table 5. Overall accuracy [%] for tree-based classifiers

Data set	Single C4.5				Ivotes / IIvotes + C4.5			
	Base	Weak	Relabel	Strong	Base	Weak	Relabel	Strong
abdominal-pain	82.84	77.45	76.87	77.92	85.20	81.77	83.21	81.30
balance-scale	78.65	73.34	72.99	73.81	84.67	80.83	80.64	79.07
breast-cancer	65.40	59.12	59.89	58.91	66.71	63.36	62.87	56.78
bupa	65.56	60.18	56.84	60.20	69.39	67.42	69.86	65.28
car	93.99	95.04	94.20	94.78	92.89	92.91	93.02	92.88
cleveland	82.25	81.52	80.98	81.86	85.08	83.83	83.70	83.70
cmc	49.25	49.27	46.58	48.46	51.57	50.69	50.98	49.45
ecoli	91.91	90.55	89.23	91.50	92.80	91.90	92.68	91.19
german	66.00	65.44	63.33	65.50	71.05	71.86	73.06	70.54
haberman	70.08	61.26	59.87	60.88	92.06	90.00	90.65	90.56
hepatits	78.47	75.93	76.16	73.74	72.55	66.67	67.39	62.88
pima	73.96	69.42	69.63	69.66	84.39	83.10	83.10	82.84
transfusion	77.75	66.15	65.61	60.85	75.65	74.14	74.24	73.23

For pairwise comparison of classifiers over all data sets we used the Wilcoxon Signed Ranks Test (confidence $\alpha = 0.05$). Considering the results of GM for tree-based classifiers (see Table 2) all single classifiers with any SPIDER preprocessing and all IIvotes ensembles were always significantly better than their baseline versions. Also all IIvotes ensembles were significantly better than single classifiers with a corresponding SPIDER option. Moreover, the IIvotes ensembles with the weak and strong options were always superior to any single classifier with any SPIDER option. After comparing pairs of Iivotes ensembles we were not able to reject the null hypothesis on equal performance for the weak and strong options, however, both of them were better than relabel.

We obtained similar results of the Wilcoxon test for rule ensembles with abstaining (see Table 4 and the left part of Table 3), although the superiority of the IIvotes ensemble with relabel over the single classifier with the same SPIDER option is slightly smaller ($p = 0.03$ while previously it was close to 0.01). Furthermore, the IIvotes ensembles with the strong option was nearly significant better than the IIvotes ensemble with the weak option ($p = 0.054$). Considering the results for the non-abstaining ensembles (Table 3), the Wilcoxon test revealed that the IIvotes ensembles weak and strong option were significantly better than the single classifiers with the same pre-processing option, however, the advantage was smaller than for the variant with abstaining.

While analysing the sensitivity alone we cannot say that IIvotes is significantly better than single classifiers with SPIDER (due to page limits we cannot show more tables with detailed results). Finally, considering the overall accuracy results of Wilcoxon test show that IIvotes integrated with SPIDER is always better than its single classifier version (see Table 5 for trees, results for rules are analogous).

5 Final Remarks

In this paper we proposed a new framework that integrates the SPIDER method for selective data pre-processing into the Ivotes ensemble. This integration aims at obtaining a better trade-off between sensitivity and specificity for the minority class than SPIDER combined with a single classifier.

Experimental results showed that the proposed IIvotes framework led to significantly better values of GM than single tree- and rule-based classifier combined with SPIDER. Despite improving the sensitivity of the minority, a satisfactory value of sensitivity is preserved, what was not achieved by SPIDER alone and other related re-sampling techniques (previous experiments [13] showed that also NCR and to some extent SMOTE suffered from decreasing specificity).

After comparing possible pre-processing options of the IIvotes framework we can say that weak and strong amplification (particularly the latter) are more efficient than relabel. Moreover, IIvotes was successful in keeping the overall accuracy at an acceptable level, comparable to baseline classifiers. Let us notice that using the standard version of Ivotes ensemble was not successful – GM did not differ significantly from values reported for single classifiers. We expect that even using a re-sampling filter to transform the whole data before constructing the ensemble is also a worse solution than integrating it inside the ensemble – see the discussion in [4].

Abstaining turned out to be a useful extension of rule ensembles as it improved their performance with respect to all considered measures. Let us remind that component classifiers in the IIvotes ensemble use unordered rule sets and the LERS classification strategy [7]. In these classifiers the conflict caused by matching a classified object to multiple rules is solved by voting with rule support. This strategy is biased toward rules from the majority classes as they are stronger and more general than rules from the minority class. This is the reason why objects from the minority class are more likely to be misclassified. Thus, refraining from making wrong predictions in some classifiers gives a chance to other component classifiers (that are more expertized for the new object) to have greater influence on the final outcome of the rule ensemble.

Our future research in processing imbalance data with rule-based ensemble classifier covers two topics. The first one is studying the impact of changing the control criterion in the ensemble from general error (or accuracy) toward measures typical for imbalanced data. The second one is exploitation of other classification strategies which could improve the role of rules for the minority class and combining them with SPIDER. This topic is a subject of our on-going research.

References

1. Blaszczynski J., Stefanowski J., Zajac M.: Ensembles of Abstaining Classifiers Based on Rule Sets. In Proc. of the 18th International Symposium on Foundations of Intelligent Systems. ISMIS2009, 2009, 382-391.
2. Breiman L.: Pasting small votes for classification in large databases and on-line. Machine Learning, 36 (1999) 85-103.
3. Chawla, N., Bowyer, K., Hall, L., Kegelmeyer, W.: SMOTE: Synthetic Minority Over-sampling Technique. J. of Artifical Intelligence Research, 16 (2002) 341-378.
4. Chawla N., Lazarevic A., Hall L., Bowyer K.: SMOTEBoost: Improving Prediction of the Minority Class in Boosting. In Proc. PKDD2003, 2003, 107–119.
5. Chawla N.: Data mining for imbalanced datasets: An overview. Chapter in Maimon O., Rokach L. (eds.): The Data Mining and Knowledge Discovery Handbook, Springer 2005, 853–867.
6. He H., Garcia E.: Learning from imbalanced data. IEEE Transactions on Data and Knowledge Engineering, vol. 21 (9), 2009, 1263–1284.
7. Grzymala-Busse J.W.: Managing uncertainty in machine learning from examples. In Proc. of the 3rd International Symposium in Intelligent Systems, 1994, 70–94.
8. Kubat, M., Matwin, S.: Addresing the curse of imbalanced training sets: one-side selection. In Proc. of the 14th Int. Conf. on Machine Learning ICML 97, (1997) 179-186.
9. Laurikkala, J.: Improving identification of difficult small classes by balancing class distribution. Tech. Report A-2001-2, University of Tampere (2001).
10. Stefanowski J.: The rough set based rule induction technique for classification problems. In Proc. of the 6th European Conf. on Intelligent Techniques and Soft-Computing EUFIT-98, 1998, 109–113.
11. Stefanowski J.: On combined classifiers, rule induction and rough sets. Transactions on Rough Sets, volume 6, 2007, 329–350.
12. Stefanowski, J., Wilk, S.: Improving Rule Based Classifiers Induced by MODLEM by Selective Pre-processing of Imbalanced Data. In Proc. of the RSKD Workshop at ECML/PKDD, Warsaw, 2007, 54–65.
13. Stefanowski J., Wilk Sz.: Selective Pre-processing of Imbalanced Data for Improving Classification Performance. In Proc. of 10th Int. Conference DaWaK 2008, LNCS vol. 5182, Springer Verlag, 2008, 283-292.