Class-dependent Dissimilarity Measures for Multiple Instance Learning

Veronika Cheplygina, David M.J. Tax, Marco Loog

Pattern Recognition Laboratory, Delft University of Technology
{v.cheplygina, d.m.j.tax, m.loog}@tudelft.nl

Abstract. Multiple Instance Learning (MIL) is concerned with learning from sets (bags) of feature vectors (instances), where the individual instance labels are ambiguous. In MIL it is often assumed that positive bags contain at least one instance from a so-called concept in instance space, whereas negative bags only contain negative instances. The classes in a MIL problem are therefore not treated in the same manner. One of the ways to classify bags in MIL problems is through the use of bag dissimilarity measures. In current dissimilarity approaches, such dissimilarity measures act on the bag as a whole and do not distinguish between positive and negative bags. In this paper we explore whether this is a reasonable approach and when and why a dissimilarity measure that is dependent on the bag label, might be more appropriate.

1 Introduction

Multiple-instance learning (MIL) [6] extends traditional supervised learning methods in order to learn from objects that are described by a set (*bag*) of feature vectors (*instances*), rather than a single feature vector only. MIL problems are often considered to be two-class problems, i.e., a bag of instances can belong either to the positive or the negative class. The bag labels are available, but the labels of the individual instances are not defined. Often assumptions are made about the instance labels and their relationship with the bag labels.

Traditional MIL problems assume that positive bags contain one or more positive instances from a so-called *concept*, whereas negative bags contain only negative instances [6, 9]. E.g. when classifying images represented by a bag of image segments as "tiger" or "no tiger", a segment containing black stripes could be seen as a positive instance for the "tiger" concept, whereas segments containing grass, sky e.t.c. would be considered negative, or background, instances.

Many traditional, "instance-based" MIL approaches try to model the concept by identifying the "most positive" instances in bags, and classify new bags as positive if they appear to have instances within this concept [6, 9]. Other, "bag-based" MIL approaches compare bags directly, using distances[18], kernels[7] or dissimilarities [17, 14]. It is possible to define a dissimilarity measure between bags, represent each bag by its dissimilarities to other bags, and use these dissimilarity values as features for supervised classifiers. A number of such

dissimilarities are investigated in [14, 17], where it is shown that some bag dissimilarities can be effective even when a concept is not clearly defined.

The instance-based methods explicitly use the assumption that positive bags are different from negative bags, whereas the bag-based methods typically do not differentiate between classes. This may not be completely natural for a MIL problem, because we have some information about how positive bags are different from negative bags. In supervised learning problems where classes are expected to behave differently, class-dependent distances[10, 5, 19] or features[2, 8] have been suggested. In this work we examine whether a similar approach might be reasonable for MIL problems.

2 Related Work

Using a class-dependent distance measure, rather than a fixed distance measure, is not a new idea. Quadratic Discriminant Analysis already allows different classes to have different covariance matrices. More attention to class-dependent distances is given in [10], where the goal is to learn weights for each feature/class combination, and to use these weights in a Mahalanobis-type metric. A similar approach is taken in [5] to improve performance in speech recognition. In [19], the authors propose learning different metrics for different classes and show that this improves classification results. In all cases, the goal is obtain high nearest neighbor performance on the learnt distances.

Other authors have examined the importance of class-dependent features rather than distances. In [2] several examples are provided where such class-dependent features are important: classification of handwritten characters, textures and documents. For instance, in a bag of words approach to document classification, it might be better to represent documents based on words that frequently occur in a particular class, as opposed to words that frequently occur in all documents. The same motivation is given in [8], where a weight is associated with each (word, class) pair. Although here, the term "dissimilarity" is used rather than distance, the learned dissimilarities are still used in a nearest neighbor setting.

For MIL, the only example of using a class-dependent dissimilarity we are aware of is from [20]. Here, bag dissimilarities are used for feature selection. The authors propose to use different dissimilarity measures for two positive bags, two negative bags, or a positive and a negative bags, to best capture the properties of the classes, such as the presence of a concept. Because the purpose is feature selection, only the dissimilarities between bags in the training set are computed. The same class-dependent dissimilarity cannot be used for the purpose of classification, because the labels of test bags are not available.

3 Review of MIL and Bag Dissimilarities

In Multiple Instance Learning, an object is represented by a bag $B_i = \{x_{ik} | k = 1, ..., n_i\} \subset \mathbb{R}^d$ of n_i feature vectors or instances. The training set $T = \{(B_i, y_i) | i =$

$1, ...N\}$ consists of positive ($y_i = +1$) and negative ($y_i = -1$) bags. The traditional assumption for MIL is that there are instance labels y_{ik} which relate to the bag labels as follows: a bag is positive if and only if it contains at least one positive instance[6]. In this case we can speak of concept (positive) instances, which are assumed to be close together in a region of the feature space called the concept $C \subset \mathbb{R}^d$, and which directly affect the bag label by their presence.

Alternatively, we can represent an object (and therefore, also a bag in a MIL problem) by its dissimilarities to prototype objects in a representation set R[11]. Often, R is taken to be the training set T, and each bag is represented as $\mathbf{d}(B_i, T) = [d(B_i, B_1), ...d(B_i, B_N)]$: a vector of dissimilarities. Therefore, each bag is represented by a single feature vector and the MIL problem can be viewed as a standard supervised learning problem.

There are various ways of defining the bag dissimilarity measure $d(B_i, B_j)$. Here we focus on defining $d(B_i, B_j)$ through the pairwise instance dissimilarities $D = [d(\mathbf{x}_{ik}, \mathbf{x}_{jl})]_{N_i \times N_j}$. We use the squared Euclidean distance for the instance dissimilarity, but other choices are also possible. In all the dissimilarities considered here, the first step is to find, for each instance in B_i, the distance to its closest instance in B_j. Using these minimum instance distances, we can define the following dissimilarities:

- Overall minimum or *minmin*: $d_{minmin}(B_i, B_j) = \min_k \min_l d(\mathbf{x}_{ik}, \mathbf{x}_{jl})$
- Average minimum or *meanmin*: $d_{meanmin}(B_i, B_j) = \frac{1}{n_i} \sum_{k=1}^{n_i} \min_l d(\mathbf{x}_{ik}, \mathbf{x}_{jl})$
- Maximum minimum or *maxmin*: $d_{maxmin}(B_i, B_j) = \max_k \min_l d(\mathbf{x}_{ik}, \mathbf{x}_{jl})$

Note that these dissimilarities are very similar to (variants) of the Hausdorff distance. However, in literature, the name "modified Hausdorff distance" has been used for a number of different distances (see [21] for some examples), so we prefer to use these more straightforward names instead. Furthermore, the Hausdorff distance is generally not symmetric, i.e. $d(B_i, B_j) \neq d(B_j, B_i)$, and often a symmetric version is obtained by taking the average or the maximum of the two values. In this paper we refrain from doing so for reasons that will become apparent in the next section.

The three dissimilarities above have their advantages and disadvantages for particular types of datasets. For instance, *minmin* performs well with a very tight concept, whereas *meanmin* is more appropriate for cases where instances from positive and negative bags arise from different distributions. A more detailed explanation is available in [4][1].

4 Class-dependent Dissimilarity

We argue that, in a MIL problem, it may be advantageous to exploit the bag label information when defining a dissimilarity between two bags. Let's assume we are dealing with a MIL problem with a well-defined concept, such as in

[1] In press, available online from http://prlab.tudelft.nl/sites/default/files/icpr2012.pdf

Figure 1(a). In this problem, if we consider all instances in a bag, any two bags may be similar or dissimilar overall. However, in MIL problems with a concept, we could speculate that the positive bags are similar at the concept level. Figure 1(b) illustrates the a dissimilarity matrix corresponding with this intuition. Each square here is a dissimilarity value between two bags, where the color of the square represents the dissimilarity value (black = 0, i.e. similar bags, white=1, i.e. dissimilar bags). Notice the difference between treating these values as distances, or as dissimilarities. In terms of distances, this representation is quite poor, because each bag has several neighbors in the opposite class. However, in terms of dissimilarities, the situation is quite different: the positive bags are clearly represented in a different way than the negative bags, so the classes are well separated.

By using the same dissimilarity to compare positive and negative bags, we risk overlooking an important difference between positive and negative bags, producing a dissimilarity matrix where all values are nearly equal. It seems that using the class information could help us capture the correct aspects of dissimilarity between bags. Ideally, we would want to have the class information of both bags when determining their dissimilarity (e.g. using the overall minimum distance for two positive bags, as in [20], but for classification purposes, it is obvious that only the labels of the prototypes are available.

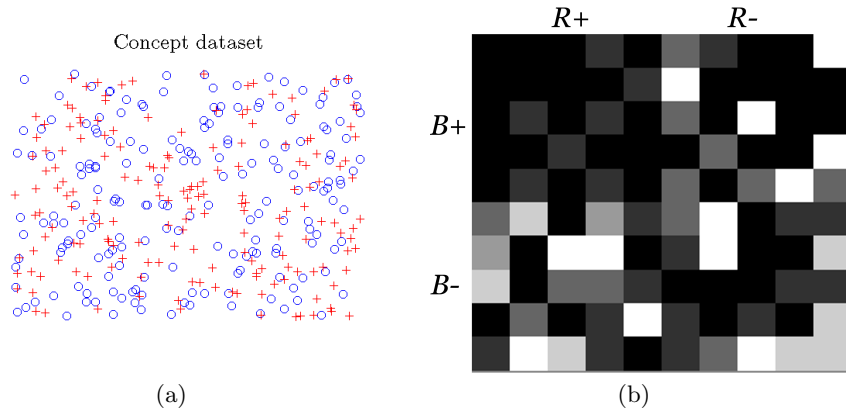

Fig. 1. (a) Artificial concept dataset, + and ◯ represent instances from positive and negative bags respectively. (b) Dissimilarity matrix reflecting the intuition we have about the positive and negative bags. The intensity value reflects the dissimilarity of two bags (black = 0, white = 1).

For positive prototypes, we want to find out something about the presence of a concept in the test bag (denoted by B), i.e. the concept instance of the prototype bag (denoted by $R+$ or $R-$ depending on the prototype label) needs to be involved. As illustrated in Figure 2, the asymmetry of the bag dissimilarities

becomes important here. If we measure the dissimilarity of the test bag to the prototype bag, denoted by $d(B \to R+)$, it may happen that none of the instances in B are matched to the positive instance in $R+$. If we measure $d(R+ \to B)$ instead, which measures the dissimilarity from the prototype to the test bag, the positive instance of $R+$ has to be matched to an instance in B. In other words, the distance from a positive prototype $d(R+ \to B)$ should be more informative. For negative prototypes, we want to highlight the absence of the concept in the prototype. In this case, we are interested in the dissimilarity to the prototype $d(B \to R-)$ because this ensures that incorrect matches (of concept instances in the test bag to background instances in the prototype) will be present.

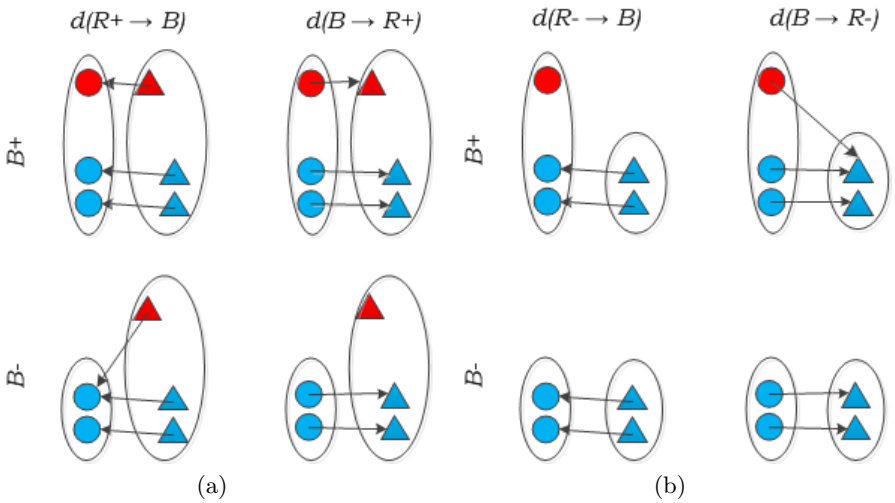

Fig. 2. Difference between the "from" $d(R \to B)$ and "to" $d(B \to R)$ dissimilarities for a prototype bag (represented by △). Each row shows the situation for a test bag (represented by ◯). The instance labels (red = positive, blue = negative) are unavailable and only shown for explanation purposes. The arrows indicate the direction of how the instance distances are measured.

4.1 Possible Dissimilarity Measure

Just the direction of measuring the dissimilarity does not yet provide us with a way to produce a single dissimilarity, but with a vector of minimum instance distances between a bag and a prototype. If the bags were very large, we could see these vectors as distributions of distances. Assuming that these distributions would be somehow different for positive and negative test bags, we could define a dissimilarity value between two bags by comparing the distributions directly.

However, in real applications, some bags may be very small (e.g. in the Musk datasets, bags with just one instance are present), so such comparisons would not

always be feasible. Instead we try to define cheap approximations for the overall bag distance, given only finite samples from the instance distance distributions. Given our previous experiences with *meanmin*, we propose to approximate both directions with the following dissimilarity:

$$d_{cd}(\cdot, R) = \begin{cases} d_{meanmin}(R \to \cdot) & \text{if } R \text{ is positive} \\ d_{meanmin}(\cdot \to R) & \text{if } R \text{ is negative} \end{cases} \quad (1)$$

5 Experiments

We test our approach on several benchmark MIL datasets:

- Musk 1, Musk 2 [6], molecule activity prediction.
- Trx Protein [16], protein function prediction.
- Mutagenesis easy, Mutagenesis hard [15], drug activity prediction.
- Fox, Tiger, Elephant [1], image classification.
- 20 Corel datasets [3], image classification.
- 25 SIVAL datasets [13], image classification.

We compare the performance of the 1-nearest neighbor classifier in the dissimilarity space using the class-dependent dissimilarity d_{cd} and its "ingredients", which we denote by d_{to} and d_{from} for brevity. In addition, we provide results of the symmetric mean $d_{avg} = \frac{1}{2}(d_{from} + d_{to})$ because before considering asymmetric dissimilarities, we have achieved good results with this symmetrized measure.

The results are given in Table 1. The class-dependent d_{cd} is performing better than d_{from} and d_{to}, which is in line with our intuition about it being able to capture more class differences. Overall, the performance of d_{avg} is comparable to that of d_{cd}, which might mean that averaging d_{from} and d_{to} captures some of the same information as in d_{cd}. For several datasets, one of these dissimilarities does significantly better than the other, although it is not entirely clear what these datasets have in common. However, it seems that in many cases where d_{to} outperforms d_{from}, d_{cd} also outperforms d_{avg} (e.g. SmileyFaceDoll).

The difference in the results of d_{from} and d_{to} is another interesting observations. For some datasets, these dissimilarities have comparable results, while for others, especially SIVAL datasets, one outperforms the other greatly. Although d_{to} is often better than d_{from}, for instance for the Apple dataset, in other datasets, such as CardboardBox, the situation is reversed.

6 Discussion and Conclusion

We have emphasized that in Multiple Instance Learning problems, it might be appropriate to treat the classes differently due to an important difference between positive and negative bags: the presence of concept instances. Most MIL approaches which compare bags directly disregard this difference. Therefore, we proposed to use a class-dependent dissimilarity based on the average minimum

Table 1. AUC performance and standard error (x100), 5x10-fold cross-validation for 1-NN classifier in the dissimilarity space. The numbers in bold indicate which dissimilarity is best (or not significantly worse than best) per dataset.

dataset	d_{to}	d_{from}	d_{cd}	d_{avg}
Musk1	**92.6 (1.1)**	**92.6 (1.1)**	**92.9 (1.2)**	**93.4 (1.1)**
Musk2	**89.7 (1.8)**	**87.2 (1.7)**	**89.7 (1.6)**	**88.5 (1.6)**
Fox	57.2 (1.4)	65.3 (1.7)	**68.7 (1.7)**	66.2 (1.7)
Tiger	**78.2 (1.6)**	**79.7 (1.3)**	75.1 (1.3)	75.6 (1.6)
Elephant	83.0 (1.3)	87.1 (1.2)	**90.8 (1.0)**	88.9 (1.0)
Protein	**62.1 (2.5)**	**61.0 (3.1)**	**63.0 (3.1)**	**64.2 (2.7)**
Mutagen easy	**89.6 (1.0)**	**89.0 (1.1)**	88.4 (1.0)	88.8 (1.0)
Mutagen hard	**79.5 (3.1)**	73.3 (3.6)	**79.7 (3.2)**	**77.3 (3.6)**
African	89.9 (0.7)	88.5 (0.7)	**90.8 (0.7)**	89.3 (0.7)
Beach	82.2 (0.8)	**82.1 (1.0)**	**83.3 (0.8)**	**82.4 (0.8)**
Historical	**87.0 (0.7)**	85.9 (0.8)	**87.1 (0.7)**	**87.6 (0.7)**
Buses	96.6 (0.3)	**97.2 (0.3)**	**97.1 (0.3)**	**96.9 (0.3)**
Dinosaurs	99.2 (0.2)	**99.7 (0.0)**	99.0 (0.2)	99.3 (0.1)
Elephants	**92.9 (0.5)**	**92.6 (0.8)**	**92.8 (0.6)**	**92.9 (0.5)**
Flowers	98.0 (0.2)	97.7 (0.3)	**98.1 (0.2)**	97.6 (0.2)
Horses	**98.9 (0.1)**	96.1 (0.4)	**98.9 (0.1)**	97.8 (0.2)
Mountains	82.7 (0.9)	82.6 (0.8)	82.5 (0.8)	**85.7 (0.7)**
Food	95.9 (0.3)	**97.0 (0.2)**	96.8 (0.3)	**97.3 (0.2)**
Dogs	84.5 (0.9)	**85.4 (0.8)**	**86.8 (0.8)**	**86.6 (0.7)**
Lizards	**93.0 (0.5)**	91.7 (0.7)	92.0 (0.6)	**92.2 (0.6)**
Fashion	90.0 (0.5)	**90.2 (0.6)**	**90.9 (0.5)**	**90.2 (0.5)**
Sunset	**94.3 (0.6)**	93.3 (0.5)	**94.1 (0.5)**	**94.7 (0.4)**
Cars	88.8 (0.7)	87.9 (0.7)	**89.8 (0.6)**	88.2 (0.6)
Waterfalls	**93.8 (0.4)**	91.1 (0.6)	93.5 (0.4)	**93.4 (0.4)**
Antique	**92.8 (0.7)**	**93.7 (0.5)**	**93.2 (0.6)**	**93.4 (0.6)**
Battleships	92.7 (0.5)	92.8 (0.4)	93.9 (0.4)	**94.5 (0.4)**
Skiing	87.0 (0.8)	**91.4 (0.6)**	87.3 (0.7)	89.9 (0.6)
Desserts	**72.0 (1.4)**	68.6 (1.1)	71.0 (1.4)	**72.7 (1.0)**
AjaxOrange	81.6 (1.5)	**85.2 (1.5)**	**86.7 (1.3)**	**86.7 (1.1)**
Apple	**69.3 (1.3)**	62.4 (1.7)	**68.6 (1.3)**	66.3 (1.5)
Banana	65.4 (1.4)	61.5 (1.7)	**66.6 (1.7)**	**65.4 (1.8)**
BlueScrunge	72.3 (1.5)	**81.2 (1.2)**	76.2 (1.4)	**81.6 (1.2)**
CandleWithHolder	80.6 (1.4)	79.7 (1.3)	85.4 (1.2)	**87.1 (1.0)**
CardboardBox	72.3 (1.5)	83.4 (1.0)	76.7 (1.4)	**86.4 (1.1)**
CheckeredScarf	95.1 (0.4)	94.2 (0.3)	95.7 (0.4)	**96.7 (0.3)**
CokeCan	85.0 (1.2)	81.8 (1.2)	**87.9 (1.1)**	**88.5 (1.1)**
DataMiningBook	84.7 (1.1)	78.6 (1.3)	**86.9 (1.1)**	**85.7 (1.2)**
DirtyRunShoes	90.9 (0.9)	90.6 (0.9)	**92.2 (0.8)**	**91.6 (0.9)**
DirtyWorkGloves	75.5 (1.6)	75.9 (1.5)	78.5 (1.5)	**82.8 (1.4)**
FabricSoftener	89.0 (1.1)	88.9 (1.1)	**95.7 (0.7)**	89.7 (1.0)
FeltFlowerRug	83.7 (1.4)	86.3 (0.9)	88.4 (1.1)	**90.2 (0.8)**
GlazedWoodPot	58.7 (1.2)	63.5 (1.7)	59.9 (1.3)	**68.1 (1.6)**
GoldMedal	75.6 (1.5)	75.1 (1.3)	80.8 (1.2)	**83.5 (1.1)**
GreenTeaBox	81.8 (1.3)	79.6 (1.4)	**85.9 (1.1)**	83.4 (1.2)
JuliesPot	68.1 (1.5)	60.3 (1.7)	**70.9 (1.5)**	64.7 (1.4)
LargeSpoon	79.0 (1.4)	71.1 (1.7)	**82.7 (1.2)**	**81.7 (1.5)**
RapBook	70.3 (1.5)	71.4 (1.5)	71.5 (1.4)	**74.1 (1.4)**
SmileyFaceDoll	**79.0 (1.4)**	59.6 (2.0)	78.2 (1.4)	75.7 (1.4)
SpriteCan	77.7 (1.2)	71.0 (1.5)	**79.5 (1.2)**	**80.2 (1.2)**
StripedNotebook	76.6 (1.2)	**78.1 (1.6)**	**78.8 (1.2)**	**77.8 (1.2)**
TranslucentBowl	67.4 (1.4)	62.1 (1.4)	**68.8 (1.5)**	63.8 (1.5)
WD40Can	86.3 (1.3)	85.6 (1.1)	**90.3 (1.0)**	89.0 (1.1)
WoodRollingPin	78.8 (1.4)	79.2 (1.3)	**82.6 (1.4)**	**82.2 (1.2)**

instance distance, which adapts itself based on the labels of the prototype bags. Experimental results showed that this class-dependent dissimilarity is indeed

more informative than the independent versions, and that it is comparable to averaging of these two dissimilarities.

In several datasets, we have noticed large differences between measuring dissimilarities from bags to prototypes (d_{to}), or from prototypes (d_{from}) to the test bags. We believe these differences may be related to the class imbalance in the Corel and SIVAL datasets, where only 4 to 5% of the bags are positive. Therefore, the representation d_{to} is actually very similar to the class-dependent representation d_{cd}. This also explains why the successes of d_{to} and d_{from} are related. In fact, the correlation coefficient between the difference of performances of d_{to} and d_{from}, and the difference of performances of d_{cd} and d_{avg}, is equal to 0.55. This suggests that d_{from} contains some information which negatively affects the performance of d_{avg}, but which can be avoided when using the dissimilarities in a class-dependent manner.

To better understand the obtained results, we also examined the performances of the individual "to" and "from" dissimilarities using only the positive, or only the negative bags as prototypes. The results were surprising, because the performances were comparable to the dissimilarities where prototypes of both classes are available. In about half of the datasets, the dissimilarity from positive prototypes outperformed all the dissimilarities in Table 1. This provides opportunities for investigating how prototype selection [12] or assigning weights to the prototype classes can further improve performance. Furthermore, this result might be of interest in MIL problems with class imbalance such as in medical image diagnosis, and is worth investigating further.

References

1. Andrews, S., Hofmann, T., Tsochantaridis, I.: Multiple instance learning with generalized support vector machines. In: Proc. of the National Conference on Artificial Intelligence. pp. 943–944. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999 (2002)
2. Bailey, A.: Class-dependent features and multicategory classification. Ph.D. thesis, Citeseer (2001)
3. Chen, Y., Bi, J., Wang, J.: Miles: Multiple-instance learning via embedded instance selection. Pattern Analysis and Machine Intelligence, IEEE Transactions on 28(12), 1931–1947 (2006)
4. Cheplygina, V., Tax, D., Loog, M.: Does one rotten apple spoil the whole barrel? In: International Conference on Pattern Recognition (In press)
5. De Wachter, M., Demuynck, K., Wambacq, P., Van Compernolle, D.: A locally weighted distance measure for example based speech recognition. In: International Conference on Acoustics, Speech, and Signal Processing. vol. 1, pp. I–181. IEEE (2004)
6. Dietterich, T., Lathrop, R., Lozano-Pérez, T.: Solving the multiple instance problem with axis-parallel rectangles. Artificial Intelligence 89(1-2), 31–71 (1997)
7. Gärtner, T., Flach, P., Kowalczyk, A., Smola, A.: Multi-instance kernels. In: Proc. of the 19th Int. Conf. on Machine Learning. pp. 179–186 (2002)
8. Kummamuru, K., Krishnapuram, R., Agrawal, R.: On learning asymmetric dissimilarity measures. In: International Conference on Data Mining. pp. 4–pp. IEEE (2005)

9. Maron, O., Lozano-Pérez, T.: A framework for multiple-instance learning. In: Advances in neural information processing systems. pp. 570–576. Morgan Kaufmann Publishers (1998)
10. Paredes, R., Vidal, E.: A class-dependent weighted dissimilarity measure for nearest neighbor classification problems. Pattern Recognition Letters 21(12), 1027–1036 (2000)
11. Pękalska, E., Duin, R.: The dissimilarity representation for pattern recognition: foundations and applications, vol. 64. World Scientific Pub Co Inc (2005)
12. Pękalska, E., Duin, R., Paclík, P.: Prototype selection for dissimilarity-based classifiers. Pattern Recognition 39(2), 189–208 (2006)
13. Rahmani, R., Goldman, S., Zhang, H., Krettek, J., Fritts, J.: Localized content based image retrieval. In: Proc. of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval. pp. 227–236. ACM (2005)
14. Sørensen, L., Loog, M., Tax, D., Lee, W., de Bruijne, M., Duin, R.: Dissimilarity-based multiple instance learning. Structural, Syntactic, and Statistical Pattern Recognition pp. 129–138 (2010)
15. Srinivasan, A., Muggleton, S., King, R.: Comparing the use of background knowledge by inductive logic programming systems. In: Proceedings of the 5th International Workshop on Inductive Logic Programming. pp. 199–230 (1995)
16. Tao, Q., Scott, S., Vinodchandran, N., Osugi, T.: Svm-based generalized multiple-instance learning via approximate box counting. In: Proc. of the 21st Int. Conf. on Machine learning. p. 101. ACM (2004)
17. Tax, D., Loog, M., Duin, R., Cheplygina, V., Lee, W.: Bag dissimilarities for multiple instance learning. Similarity-Based Pattern Recognition pp. 222–234 (2011)
18. Wang, J.: Solving the multiple-instance problem: A lazy learning approach. In: Proc. of the 17th Int. Conf. on Machine Learning (2000)
19. Weinberger, K., Saul, L.: Distance metric learning for large margin nearest neighbor classification. The Journal of Machine Learning Research 10, 207–244 (2009)
20. Zafra, A., Pechenizkiy, M., Ventura, S.: Reducing dimensionality in multiple instance learning with a filter method. Hybrid Artificial Intelligence Systems pp. 35–44 (2010)
21. Zhao, C., Shi, W., Deng, Y.: A new hausdorff distance for image matching. Pattern Recognition Letters 26(5), 581–586 (2005)