Risk Minimization Framework for Multiple Instance Learning from Positive and Unlabeled Bags

Han Bao
The University of Tokyo,
113-0033 Tokyo, Japan
tsutsumi@ms.k.u-tokyo.ac.jp

Tomoya Sakai
The University of Tokyo,
277-8561 Chiba, Japan
RIKEN Center for AIP,
103-0027 Tokyo, Japan
sakai@ms.k.u-tokyo.ac.jp

Masashi Sugiyama
RIKEN Center for AIP,
103-0027 Tokyo, Japan
The University of Tokyo,
277-8561 Chiba, Japan
sugi@k.u-tokyo.ac.jp

Issei Sato
The University of Tokyo,
277-8561 Chiba, Japan
RIKEN Center for AIP,
103-0027 Tokyo, Japan
sato@k.u-tokyo.ac.jp

Multiple instance learning (MIL) is a variation of traditional supervised learning problems with ambiguous label information. In standard MIL setting, a classifier is trained from *positive bags* which are sets of instances with at least one positive instance, and *negative bags* which are sets of only negative instances. This problem setting originally came from a biochemistry problem (Dietterich, Lathrop, and Lozano-Pérez, 1997), and has also been used in the fields of automatic image annotation and text categorization (Andrews, Tsochantaridis, and Hofmann, 2002).

Previous work often assumes training bags are fully labeled. Yet, it is often difficult to obtain an enough amount of labeled data while many unlabeled data are often available in practical situations such as outlier detection. In the framework of PU (positive and unlabeled) learning (du Plessis, Niu, and Sugiyama, 2015), a classifier can be trained only from positive and unlabeled data, which reduces labeling costs. Wu et al. (2014) proposed an algorithm of multiple instance learning from positive and unlabeled bags (PU-MIL) that imputes labels to unlabeled bags at random and refines the performance of the classifier by the genetic algorithm. However, this method remains to be non-convex, which is computationally inefficient to optimize.

In this work, we propose a novel method of PU-MIL based on the convex PU learning framework (du Plessis, Niu, and Sugiyama, 2015) together with the set kernels (Gärtner et al., 2002). The set kernels map multiple instances (sets of instances) to a feature space. Thus, they can be treated by traditional instance-level methods. Relationships between PU-MIL and the related problems are shown in Figure 1. We experimentally evaluate the proposed method on synthetic and benchmark dataset on biochemistry and image annotation tasks[1] and show the proposed method outperforms the existing method[2]. Our proposed method is also shown to be more computationally efficient than the existing method.

More details are available in our arXiv paper[3].

[1] http://www.cs.columbia.edu/~andrews/mil/datasets.html

[2] The classification performances are evaluated by AUC since class-imbalance case is assumed.

[3] https://arxiv.org/abs/1704.06767

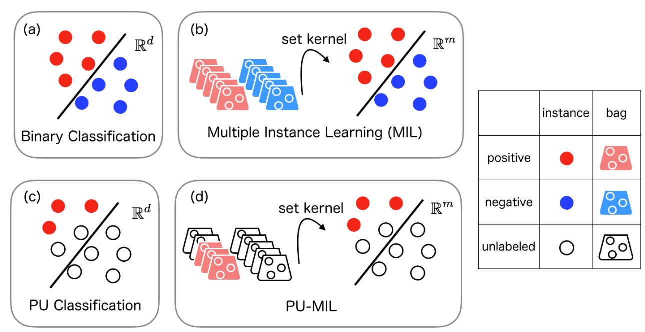

Figure 1: Schematics of PU-MIL and related problems. In this work, we consider (d) multiple instance learning from positive and unlabeled bags (PU-MIL), combining the set kernels (Gärtner et al., 2002) and PU classification (du Plessis, Niu, and Sugiyama, 2015).

References

Andrews, S.; Tsochantaridis, I.; and Hofmann, T. 2002. Support vector machines for multiple-instance learning. In *Advances in Neural Information Processing Systems 15*, 577–584. Cambridge, MA, USA: MIT Press.

Dietterich, T. G.; Lathrop, R. H.; and Lozano-Pérez, T. 1997. Solving the multiple instance problem with axis-parallel rectangles. *Artificial Intelligence* 89(1):31 – 71.

du Plessis, M. C.; Niu, G.; and Sugiyama, M. 2015. Convex formulation for learning from positive and unlabeled data. In *Proceedings of the 32nd International Conference on Machine Learning (ICML-15)*, 1386–1394. JMLR Workshop and Conference Proceedings.

Gärtner, T.; Flach, P. A.; Kowalczyk, A.; and Smola, A. J. 2002. Multi-instance kernels. In *In Proc. 19th International Conf. on Machine Learning*, 179–186. Morgan Kaufmann.

Wu, J.; Zhu, X.; Zhang, C.; and Cai, Z. 2014. Multi-instance learning from positive and unlabeled bags. In *Proceeding of the 18th Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 237–248. Cham: Springer International Publishing.