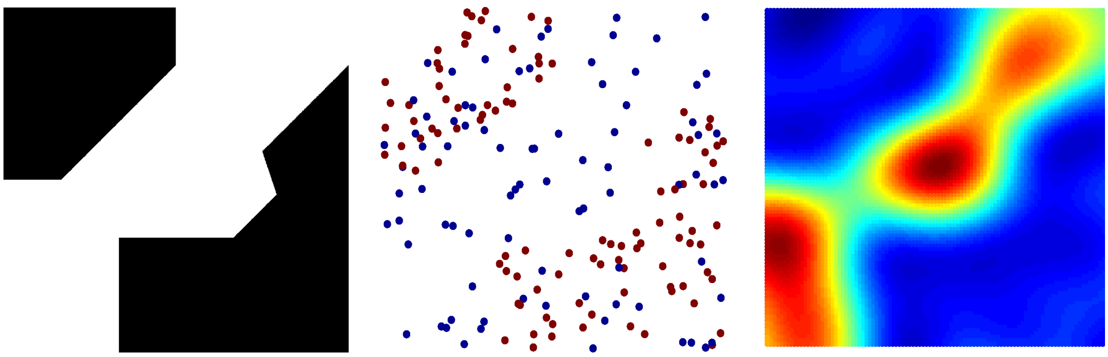
Convex Multiple-Instance Learning

Fuxin Li and Cristian Sminchisescu
Computer Vision and Machine Learning Group, INS, University of Bonn

Multiple-Instance learning has long been considered a hard non-convex problem, and attacked using sophisticated non-convex methods like alternative minimization, deterministic annealing [1] or convex-concave procedures [2]. In this work we pursue instead a two-step approach that solves a novel multiple instance learning formulation based on convex optimization. In the first step, we jointly estimate a score for each instance and a regressor that fits this scoring. In the second step, the scores are used to build a feature vector and classification is achieved using a simple linear SVM. Using this approach, we could perform accurate bag classification, decide instance labels and provide a ranking based on the relevance of the instances. We also propose an estimation method to assess the percentage of true positive instances (*witnesses*) inside a positive bag that is shown to be fairly accurate. Experimental results in synthetic and benchmark datasets demonstrate the encouraging performance of this approach.

First we perform an experiment on the synthetic dataset used in Gehler and Chapelle [1], a 2-D dataset with the actual decision boundary shown in fig. 1 (a). This controlled setting helps us to understand the algorithm's behavior. The positive bags have a fraction of points sampled uniformly from the positive region (in white) and the rest sampled uniformly from the negative region (in black). An example at 40% witness rate is shown in fig. 1 (b). The plotted labels of the instances are labels of their bags, therefore one could see many positive (blue) instances in the negative (red) region. To test the effect of witness rates on the algorithm, ten different types of datasets are created by varying the witness rates over $0.1, 0.2, \ldots, 1$. We show an example of a scoring function trained on fig. 1(b) in in fig. 1 (c). With this scoring function, accurate classification in this dataset is straightforward: a simple thresholding rule is necessary to achieve near perfect classification. This illustrates the main strength of the approach: it can learn exactly, using convex optimization, a scoring function that makes classification easy.



(a) (b) (c)

Figure 1: Synthetic Dataset (Best viewed in color). (a) The true decision boundary. (b) Training points at 40% witness rate. (c) The learned regression function.

Topic: Machine Learning, Multiple Instance Learning **Preference: poster**

References

[1] Peter Gehler and Olivier Chapelle. Deterministic annealing for multiple-instance learning. In Meila, M., and X. Shen, editors, *Proceedings of the 11th International Conference on Artificial Intelligence and Statistics*, pages 123–130, March 2007.

[2] Pak ming Cheung and James T. Kwok. A regularization framework for multiple-instance learning. In *In Proceedings of the Twenty-Third International Conference on Machine Learning*, pages 193–200. ACM Press, 2006.