Support Vector machine for multiple instance learning

Stuart Andrews
stu@cs.brown.edu
Ioannis Tsochantridis
it@cs.brown.edu
Thomas Hofmann
th@cs.brown.edu

Department of Computer Science,
Brown University

This paper proposed how we can use maximum marginal property of SVM to solve problems in manner which is accuracy is getting better by implementing on pharmaceutical data set . Multiple-instance learning (MIL) is a type of supervised learning. Instead of receiving a set of instances which are individually labeled, the learner receives a set of labeled bags, each containing many instances.SVM is a margin based classifier having excellent generalization capabilities. Such models try to find an optimal separating hyper plane between data points of different classes in high dimensional space.

The mixed formulations of MIL as a generalized soft margin SVM can be written as follow in primal form

$$minmin\frac{1}{2}||w^2||+C\sum\alpha i$$

The mi-SVM formulation leads to a mixed integer programming problem. One has to find both the optimal labeling and the optimal hyperplane. On a conceptual level this mixed integer formulation captures exactly what MIL is about, i.e. to recover the unobserved pattern labels and to simultaneously find an optimal discriminant.Yet, this poses a computational challenge since the resulting mixed integer programming problem cannot be solved efficiently with state-of-the-art tools, even for moderate size data sets.

we can use this method in image processing for annotating image.we can also use this for classification of drugs on the basis of molecule size .

then comes the part of optimization.A general scheme for a simple optimization heuristic may be described as follows. Alternate the following two steps: (i) for given integer variables, solve the associated QP and find the optimal discriminant function, (ii) for a given discriminant, update one, several, or all integer variables in a way that (locally) minimizes the objective. The latter step may involve the update of a label variable Yi of a single pattern in miSVM, the update of a single selector variable 8(I) in MI-SVM, or the simultaneous update of all integer variables. Since the integer variables are essentially decoupled given the discriminant (with the exception of the bag constraints in mi-SVM), this can be done very efficiently. Also notice that we can re-initialize the QP-solver at every iteration with the previously found solution, which will usually result in a significant speed-up. In terms of initialization of the optimization procedure, we suggest to impute positive labels for patterns in positive bags as the initial configuration in mi-SVM. In MI-SVM, XI is initialized as the centroid of the bag patterns. Figure 2 and 3 summarize pseudo-code descriptions for the algorithms utilized in the experiments.

An alternative way of applying maximum margin ideas to the MIL setting is to extend the notion of a margin from individual patterns to sets of patterns. It is natural to define the functional margin of a bag with respect to a hyperplane by $\gamma i == \mathrm{YI} max((w, Xi)+b)$.

for a positive bag the margin is defined by the margin of the "most positive" pattern, while the margin of a negative bag is defined by the "least negative" pattern

After experimenting this method on musk data-set ,image data-set , TREC9 We have presented a novel approach to multiple-instance learning based on two alternative generalizations of the maximum margin idea used in SVM classification. Although these formulations lead to hard mixed integer problems, even simple local optimization heuristics already yield quite competitive results compared to the baseline approach. This method is totally based on herusitcal approch so there will be some uncertainty in results but as far the process the maximization will be the same every time because the iteration will always stoped at local minima .

Reference

P. Auer. On learning from multi-instance examples: Empirical evaluation of a theoretical approach. In Proc. 14th International Conf. on Machine Learning, pages 21- 29. Morgan Kaufmann, San Francisco, CA, 1997.

C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik. Blobworld: A system for region-based image indexing and retrieval. In Proceedings Third International Conference on Visual Information Systems. Springer, 1999.

A. Demirez and K. Bennett. Optimization approaches to semisupervised learning. In M. Ferris, O. Mangasarian, and J. Pang, editors, Applications and Algorithms of Complementarity. Kluwer Academic Publishers, Boston, 2000.

T . G. Dietterich, R. H. Lathrop, and T . Lozano-Perez. Solving the multiple instance problem with axis-parallel rectangles. Artificial Intelligence, 89(1-2):31- 71 , 1997.

T. Gartner, P. A. Flach, A. Kowalczyk, and A. J. Smola. Multi-instance kernels. In Proc. 19th International Conf. on Machine Learning. Morgan Kaufmann, San Francisco, CA, 2002.

T. Joachims. Transductive inference for text classification using support vector machines. In Proceedings 16th International Conference on Machine Learning, pages 200- 209. Morgan Kaufmann, San Francisco, CA, 1999.

P.M. Long and L. Tan. PAC learning axis aligned rectangles with respect to product distributions from multiple-instance examples. In Proc. Compo Learning Theory, 1996.

O. Maron and T. Lozano-Perez. A framework for multiple-instance learning. In Advances in Neural Information Processing Systems, volume 10. MIT Press, 1998.

O. Maron and A. L. Ratan. Multiple-instance learning for natural scene classification. In Proc. 15th International Conf. on Machine Learning, pages 341- 349. Morgan Kaufmann, San Francisco, CA, 1998.

J. Ramon and L. De Raedt. Multi instance neural networks. In Proceedings of ICML2000, Workshop on Attribute- Valu e and Relational Learning, 2000.

B. SchOlkopf and A. Smola. Learning with Kernels. Support Vector Machines, Regularization, Optimization and Beyond. MIT Press, 2002.