Analysis of the paper

November 6, 2021

Yawer Abbas
B20CI050

Critical Analysis

MIL is supervised learning where you need labeled classes to classify raw data into small bags. For example, if a bag is positive in at least one of its member patterns, it is considered part of a positive bag. Otherwise, the bag will go into a negative bag. If you categorize all of these bags, you will see some patterns in these positive and negative bags, which will be used for further classification. It is a pattern level classifier, so it is different from other classifiers. Introduced two ways to apply SVM to MIL related issues. The first approach is to explicitly treat each pocket as chaotic, like an integer label, and then maximize the margin or soft margin for hidden labels. The second approach is to first identify the bag and then maximize the margin for the bag. A better option as the first approach is a kind of busy learning and the second approach is lazy learning so there are predefined pockets and the rest just puts the new data in the predefined classes. is. Note that the standard classification uses the label $y.Assign$ i according to x i and an unknown number if the data does not belong to a negative bag. Unlabeled data is now used as the decision boundary by maximizing the margins at all data points. Each unlabeled pattern can be labeled individually with inductive inference, but the pattern labeling in the MIL positive bag is combined by an inequality constraint. Repeat the hyperplane and labeling system until optimal conditions are reached. Note that in miSVM, the edge of each instance is important and you can set the instance label variable under bag label constraints to maximize the edge. By comparison, in MISVM, only one instance per bag is important because it determines the bag margin. The former is suitable for tasks where the user cares about instance labels, and the latter is suitable for tasks where only pocket labels are concerned. Both methods were implemented using mixed integer quadratic programming. Implemented this model in MUSK data, image data, and text classification data. The MUSK dataset is the benchmark dataset used by almost all approaches to date and described in detail in the Landmark Paper [4]. The datasets MUSK1 and MUSK2 consist of a description of molecules with multiple low energy conformations. MiSVM delivers a competitive level of accuracy on both

MUSK1 and -MUSK2 records. MISVM is superior to miSVM in MUSK2, but significantly inferior in MUSK1. Both methods do not perform as well as (repetitive APR) 2, but is comparable to other approaches to MIL. The fresh image data is used to run miSVM. The original data is a color image of the Corel dataset preprocessed by the Blobworld system and segmented into [2]. In this figure, image consists of a set of segments (or blobs), each segment identified by a descriptor of color, texture, and shape . The experiment used 3 different categories (”elephant”, ”fox”, ”tiger”) . Each dataset contains 100 positive and 100 negative sample images. The latter was randomly selected from a pool of photographs of 4,444 other animals. Due to the limited accuracy of image segmentation, the relatively small number of area descriptors, and the small size of the training set, this is a fairly difficult classification problem. The MISVM performed slightly better than the miSVM, making both heuristic methods more susceptible to other local local minima values. Generated a MIL record for text classification. It is based on the publicly available TREC9 dataset (also known as OHSUMED). This data provides an interesting additional benchmark, as the display is very sparse and high-dimensional compared to other datasets. For text classification tasks, the significant difference between the two methods is not clear.

Pros/Cons

Pros:

1. SVM works better at larger sizes, MIL uses multiple labels, so it is less complex and gives better results. Creating preinstalled packages.

2. Since the proposed method is simple to implement and can perform another downsizing property to reduce storage space, it is likely to be used for drug recognition that impedes data storage.

Disadvantages:

1. SVM cannot be implemented on nonlinear data.

2. Result of multiple integers problem was not effectively solved.

3. Optimization of the SVM parameters has a significant impact on both the training time and the accuracy of the wells of the decoding model.

Future work

On the issue of MIL research, we looked at a wider range of records and usages than usual, and obtained very good results of on various records. I strongly doubt that many MIL methods have been optimized to work well with the MUSK benchmark. The dataset used in the experiment will be open to the public to facilitate further empirical comparisons.

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.