Self-Training with Adaptive Regularization for S^{3}VM

Edward Cheung   Yuying Li

Abstract

The Semi-Supervised Support Vector Machine (S^{3}VM) solves a non-convex, Mixed-Integer Program (MIP). Due to the difficulty of the problem, convex approximations have typically been used. However, existing approaches have suffered from poor scalability and struggle on *manifold-like* datasets. The poor predictive performance suggests that for some datasets, convex approximations may not be a sufficiently accurate approximation to the problem. We present a self-training approach with self-adapting regularization parameters for S^{3}VM formulations. At each iteration, the regularization parameters are adapted to better reflect label confidence, class proportion, and to gradually include more unlabeled points. We show that updating the S^{3}VM framework iteratively, the proposed method can better capture the manifold assumption and achieve strong performance where S^{3}VM methods typically struggle on benchmark datasets.