PULNS: Positive-Unlabeled Learning with Effective Negative Sample Selector

Abstract: Positive-unlabeled learning (PU learning) is an important case of binary classification where the training data only contains positive and unlabeled samples. The current state-of-the-art approach for PU learning is the cost-sensitive approach, which casts PU learning as a cost-sensitive classification problem and relies on unbiased risk estimator for correcting the bias introduced by the unlabeled samples. However, this approach requires the knowledge of class prior and is subject to the potential label noise. In this paper, we propose a novel PU learning approach dubbed PULNS, equipped with an effective negative sample selector, which is optimized by reinforcement learning. Our PULNS approach employs an effective negative sample selector as the agent responsible for selecting negative samples from the unlabeled data. While the generated negative samples can be used to improve the classifier, the performance of classifier is also used as the reward to improve the selector through the REINFORCE algorithm. By alternating the updates of the selector and classifier, the performance of both are improved. Extensive experimental studies on 7 real-world application benchmarks also demonstrate that PULNS consistently outperforms the current state-of-the-art methods in PU learning, and our experimental results also confirm the effectiveness of the negative sample selector underlying PULNS.

Camera ready version to come soon.