

io-port 06378120**Agarwal, Shivani****Surrogate regret bounds for bipartite ranking via strongly proper losses.**

J. Mach. Learn. Res. 15, 1653-1674 (2014).

Summary: The problem of bipartite ranking, where instances are labeled positive or negative and the goal is to learn a scoring function that minimizes the probability of mis-ranking a pair of positive and negative instances (or equivalently, that maximizes the area under the ROC curve), has been widely studied in recent years. A dominant theoretical and algorithmic framework for the problem has been to reduce bipartite ranking to pairwise classification; in particular, it is well known that the bipartite ranking regret can be formulated as a pairwise classification regret, which in turn can be upper bounded using usual regret bounds for classification problems. Recently, Kotlowski et al. showed regret bounds for bipartite ranking in terms of the regret associated with balanced versions of the standard (non- pairwise) logistic and exponential losses. In this paper, we show that such (non-pairwise) surrogate regret bounds for bipartite ranking can be obtained in terms of a broad class of proper (composite) losses that we term as *strongly proper*. Our proof technique is much simpler than that of Kotlowski et al., and relies on properties of proper (composite) losses as elucidated recently by Reid and Williamson and others. Our result yields explicit surrogate bounds (with no hidden balancing terms) in terms of a variety of strongly proper losses, including for example logistic, exponential, squared and squared hinge losses as special cases. An important consequence is that standard algorithms minimizing a (non-pairwise) strongly proper loss, such as logistic regression and boosting algorithms (assuming a universal function class and appropriate regularization), are in fact consistent for bipartite ranking; moreover, our results allow us to quantify the bipartite ranking regret in terms of the corresponding surrogate regret. We also obtain tighter surrogate bounds under certain low-noise conditions via a recent result of Clemencon and Robbiano.

Keywords: ranking; area under ROC curve (AUC); statistical consistency; regret bounds; proper losses; strongly proper losses

<http://jmlr.csail.mit.edu/papers/v15/agarwal14b.html>