Volume Regularization for Binary Classification

Koby Crammer

Tal Wagner
Volume Regularization for Binary Classification

Koby Crammer    Tal Wagner    [Tu46]

Results on binary 30 NLP datasets

Compared with - SVM - AROW (online learning with adaptive weights)

Points above diagonal line indicates superiority of BoW
Volume Regularization for Binary Classification

Koby Crammer    Tal Wagner    [Tu46]

Number of wins of BoW vs SVM on all 45 pairs of USPS data-set
Empirical Risk Minimization with Regularization

\[ w^* = \arg \min_w \left( \frac{1}{n} \sum_i \ell(w, (x_i, y_i)) + \text{const} \mathcal{R}(w) \right) \]

Same, over boxes of weights ...

\[ B^* = \arg \min_B \sup_{w \in B} \left( \frac{1}{n} \sum_i \ell(w, (x_i, y_i)) + \text{const} \mathcal{R}(w) \right) \]

Two batch formulations

Optimization algorithm guaranteed to converge

Generalization bound based on PAC-Bayes

Experiments …