Machine Learning: Boosting

October 10, 2015 Jordan Boyd-Graber

Problems from *Foundations of Machine Learning* by Mohri et al.

1. Simplified AdaBoost. Suppose we simplify AdaBoost by setting the parambeter α_t to a fixed value $\alpha_t = \alpha > 0$, independent of round t.

   (a) Let γ be such that $\left(\frac{1}{2} - \epsilon_t\right) \geq \gamma > 0$. Find the optimal α as a function of γ wrt empirical error.

   (b) For the value of α that you found, does the algorithm assign the same probability mass to correctly classified and misclassified examples at each round? If not, which set has higher probability?

   (c) Using the previous value of α, give a bound on the empirical error of the algorithm as a function only of γ and the number of rounds of boosting (T).

2. Update guarantee. Assume that the main weak learner assumption of AdaBoost holds. Let h_t be the base learner selected at round t. Show that the base learner h_{t+1} selected at round $t + 1$ must be different from h_t.

3. Fix $\epsilon \in (0, .5)$. Let the training data be m points. $\frac{m}{4}$ negative points are each at $(1, 1)$ and $(-1, -1)$; $\frac{m(1-\epsilon)}{4}$ points are each at $(1, -1)$; and $\frac{m(1+\epsilon)}{4}$ points are at $(-1, +1)$. What does AdaBoost do with boosting stumps as the base learner? What solution does the algorithm return after T rounds?