Thomas Kesselheim

January 22, 2019
Due: lecture on January 29, 2019

# Algorithms and Uncertainty

Winter Semester 2018/19

Exercise Set 11

**Exercise 1:** (4 Points)
Consider a set of data points in $\mathbb{R}$ with arbitrary binary labels $S = \{(x_1, y_1), \dots, (x_m, y_m)\} \subseteq \mathbb{R} \times \{-1, 1\}$. Construct a hypothesis $h^*$ as a linear combination of decision stumps such that $h(x_i) = y_i$ for all $i$.

**Exercise 2:** (4 Points)
Show that for all steps $t$ in AdaBoost we have

$$\sum_{i:h_t(x_i)\neq y_i} p_i^{(t+1)} = \frac{1}{2} \ .$$

This also gives a nice intuitive interpretation of the choice of $p^{(t+1)}$: It is a probability distribution such that hypothesis $h_t$ is not better than random guessing.

**Exercise 3:** (3+4 Points)
We consider two modified versions of the Perceptron algorithm.

(a) Replace the update step by $\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} + \eta y_i \mathbf{x}_i$ for some $\eta > 0$. Show that this algorithm takes the same number of iterations and computes a vector in the same direction as the original Perceptron algorithm.

(b) Start from any vector $\mathbf{w}^{(1)}$. Give a bound on the number of iterations depending on $R$, $B$, and $\|\mathbf{w}^{(1)}\|_2$.

**Exercise 4:** (4+1 Points)
Consider $X = \mathbb{R}$, $Y = \{-1, 1\}$, $\mathcal{H}$ as the set of decision stumps, and 0/1 loss $\ell(h, z) = 0$ if $h$ classifies $z$ correctly and 1 otherwise.

(a) Show that Empirical Risk Minimization is not universally $\delta$-replace-one stable for any $\delta < 1$, even for arbitrary sizes of $S$. That is, for every $m \in \mathbb{N}$ give $S$, $i$, and $z'$ such that $\ell(h_{S^i}, z_i) - \ell(h_S, z_i) \geq 1$.

(b) In which lecture did we show that the generalization error still vanishes for $m \to \infty$ and why?