University of Salzburg | Lecturer: Roland Kwitt

Machine Learning (911.236)

Exercise sheet **B**

Exercise 1. 4 P.

Let $\mathcal{H}$ be a hypothesis class of binary classifiers. Show that if $\mathcal{H}$ is agnostic PAC learnable, then $\mathcal{H}$ is PAC learnable as well. Furthermore, if A is a successful agnostic PAC learner for $\mathcal{H}$, then $\mathcal{H}$ is also a successful PAC learner for $\mathcal{H}$.

Exercise 2. 6 P.

Suppose you have an algorithm, A, such that if the size m of the training set $S = ((x_1, y_1), \dots, (x_m, y_m))$ satisfies

$$m \geq m_{\mathcal{H}}(\epsilon)$$

for $\epsilon \in (0, 1)$, it holds that for any distribution $\mathcal{D}$ over $\mathcal{X} \times \{0, 1\}$, we have

$$\mathbb{E}_{S \sim \mathcal{D}^m}[L_{\mathcal{D}}(A(S))] \leq \min_{h'} L_{\mathcal{H}}(h') + \epsilon \ \ .$$

Your task is to show that for every $\delta \in (0, 1)$, if

$$m \geq m_{\mathcal{H}}(\delta\epsilon) \ \ ,$$

then with probability of (at least) $1 - \delta$, it holds that

$$L_{\mathcal{D}}(A(S)) \leq \min_{h' \in \mathcal{H}} L_{\mathcal{D}}(h') + \epsilon \ \ .$$

Hint: Use Markov's inequality at the appropriate place.

Total #points: 10 P.