Thomas Kesselheim

January 15, 2019
Due: lecture on January 22, 2019

Algorithms and Uncertainty

Winter Semester 2018/19

Exercise Set 10

Exercise 1: (5 Points)
We consider again $X = \mathbb{R}$ and let $\mathcal{H}$ be the hypotheses of the form $h(x) = 1$ for $x \in [a, b]$, $h(x) = 0$ otherwise. What is the VC dimension of $\mathcal{H}$? Prove your claim.

Exercise 2: (3+4 Points)
Let $X = \mathbb{R}^2$ and let $\mathcal{H}$ be the linear classifiers, that is, each classifier is defined by a line and a direction (above or below) such that the points above and on the line are classified as 1, the other ones as -1.

(a) Give an example that the VC dimension is at least 3.

(b) Show that the VC dimension is less than 4.

Exercise 3: (3 Points)
Let $\mathcal{H}$ and $\mathcal{H}'$ be two hypothesis classes with $\mathcal{H} \subseteq \mathcal{H}'$. Show that the VC dimension of $\mathcal{H}$ is at most as large as the one of $\mathcal{H}'$.

Exercise 4: (2 Points)
Show that if a hypothesis class is not PAC learnable in the realizable sense it is not PAC learnable in the agnostic sense either.

Exercise 5: (3 Points)
Suppose that you are given a learning algorithm, which only approximately minimizes the training error. That is, for all training sets S, we have $\mathrm{err}_S(h_S) \leq \min_{h \in \mathcal{H}} \mathrm{err}_S(h) + \gamma$ for some fixed $\gamma > 0$.
Show that for all choices of $\epsilon > 0$, $\delta > 0$, if S is a training set of size

$$m \geq \frac{8}{\epsilon^2} \ln \left(\frac{4\mathcal{H}[2m]}{\delta} \right) \quad .$$

then with probability at least $1 - \delta$, we have $\mathrm{err}_{\mathcal{D}}(h_S) \leq \min_{f \in \mathcal{H}} \mathrm{err}_{\mathcal{D}}(f) + \epsilon + \gamma$.