COLT— Exercise Sheet 1

Exercise 1. Generalize the algorithm for rectangle learning to $\mathbb{R}^n$ and prove that the class of axis-aligned hyperrectangles in $\mathbb{R}^n$ is efficiently PAC-learnable.

Exercise 2. Consider $X = \mathbb{R}^2$ and the concept class $\mathcal{C}_c$ consisting of circles centered at $(0, 0)$, i.e., sets of the form $\mathbf{c}_r = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r\}$. Show that this class is efficiently PAC-learnable and that the sample bound is $m \geq \frac{1}{\epsilon} \log(\frac{1}{\delta})$.

Exercise 3. Consider $X = \mathbb{R}$. Let $\mathcal{C}_I$ be the concept class of closed intervals for the form $[a, b]$, where $a, b \in \mathbb{R}$. Show that $\mathcal{C}^I$ is efficiently PAC-learnable.

Exercise 4. Consider $X = \mathbb{R}$ and $p > 1$. Let $\mathcal{C}^I_p$ be the class of unions of p (possibly overlapping) intervals $[a_1, b_1] \cup \ldots \cup [a_p, b_p]$. Show that this class is efficiently PAC-learnable. Derive a formula depending on ϵ, δ and p for the number of samples m required to learn a concept from $\mathcal{C}^I_p$.

Exercise 5. Consider $X = \bigcup_{n>0}\{0, 1\}^n$. Let $\mathcal{C}_B$ be the class of Boolean functions $\mathbf{c}\colon X \to \{0, 1\}^*$. Show that there is no efficient PAC-learning algorithm for $\mathcal{C}_B$.

Exercise 6. Consider $X = \bigcup_{n>0}\{0, 1\}^n$. Let $\mathcal{C} \subseteq \mathcal{C}_B$ be a subclass of all Boolean functions $\mathbf{c}\colon X \to \{0, 1\}$. We pick some binary representation of Boolean functions and define $size(\mathbf{c})$ as the minimal representation size of the function $\mathbf{c}$.

Assume that there is an efficient PAC-learning algorithm for $\mathcal{C}$, which additionally takes $size(c)$ as input. Show that there is also an efficient PAC-learning algorithm for $\mathcal{C}$, which does not have access to $size(c)$.

Exercise 7. Consider the following modification of the PAC framework, called the **two-oracle** variant. For every concept $\mathbf{c}$, we consider two arbitrary probability distributions D_+ and D_- over respectively positive and negative samples classified by $\mathbf{c}$. The learning algorithm has access to two random oracles: (a) the positive oracle, which returns positive (for $\mathbf{c}$) examples drawn randomly according to the distribution D_+, and (b) the negative oracle, which returns negative examples drawn according to D_-. We require that the learning algorithm invoked with parameters ϵ, δ returns a hypothesis $\mathbf{h}$ such that with probability at least $1 - \delta$ we have: (1) $\mathbb{P}_{x \sim D_+}(\{x \mid \mathbf{h}(x) = 0\}) \leq \epsilon$ (false negatives), and (2) $\mathbb{P}_{x \sim D_-}(\{x \mid \mathbf{h}(x) = 1\}) \leq \epsilon$ (false positives).

Let h_0 (resp., h_1) be the function that labels every example with 0 (resp., 1). Show that $\mathcal{C}$ is efficiently PAC-learnable using $\mathcal{H}$ in the original model if and only if $\mathcal{C}$ is efficiently PAC-learnable using $\mathcal{H} \cup \{h_0, h_1\}$ in the two-oracle model.