Exercises
Advanced Machine Learning
Fall 2019

Series 8, 16 Dec 2019 (PAC learning)

Machine Learning Laboratory
Dept. of Computer Science, ETH Zürich
Prof. Joachim M. Buhmann
Web https://ml2.inf.ethz.ch/courses/aml/

Teaching assistant: **Viktor Gal**
viktor.gal@inf.ethz.ch

Problem 1 (PAC learnability):

Let $\mathcal{C}$ be a concept class and $\mathcal{A}$ a learning algorithm that, when given a sample of n instances as input, outputs the empirical-risk minimizer $\hat{c}_n^* \in \mathcal{C}$. Consider the following inequalities:

- For any $n \in \mathbb{N}$,
$$\mathcal{R}\left(\hat{c}_n^*\right) - \inf_{c \in \mathcal{C}} \mathcal{R}\left(c\right) < 2 \sup_{c \in \mathcal{C}} \left| \hat{\mathcal{R}}_n\left(c\right) - \mathcal{R}\left(c\right) \right|.$$

- If $\mathcal{C}$ is finite, $n \in \mathbb{N}$, and $\epsilon > 0$,
$$\mathbf{P}\left(\sup_{c \in \mathcal{C}} \left| \hat{\mathcal{R}}_n\left(c\right) - \mathcal{R}\left(c\right) \right| > \epsilon \right) \leq 2\left|\mathcal{C}\right| \exp\left(-2n\epsilon^2\right).$$

Use these inequalities to prove that $\mathbf{P}\left(\mathcal{R}\left(\hat{c}_n^*\right) - \inf_{c \in \mathcal{C}} \mathcal{R}\left(c\right) > \epsilon\right) \to 0$, as $n \to \infty$.

Problem 2 (PAC learnability of intervals):

Consider the concept class $\mathcal{C} := \left\{ \mathbb{I}_{[\ell,\infty)} \mid \ell \in \mathbb{R} \right\}$ consisting of all indicator functions of intervals of the form $[\ell, \infty)$:
$$\mathbb{I}_{[\ell,\infty)}(x) = \begin{cases} 0 & x < \ell \\ 1 & x \geq \ell \end{cases}$$

Fix some $\ell^* \in \mathbb{R}$ and consider the classifier $c^* = \mathbb{I}_{[\ell^*,\infty)}$. Let $X_1, X_2, \ldots$ be i.i.d. random variables taking values in $\mathbb{R}$ and let $Y_i := c^*\left(X_i\right)$. Moreover, for $n \in \mathbb{N}$, let
$$X_{\min}^n := \min_{i \leq n, Y_i = 1} X_i \quad \text{and} \quad \hat{c}_n := \mathbb{I}_{[X_{\min}^n, \infty)}.$$

1. Let $\epsilon > 0$ and assume that there is a unique $\ell_\epsilon^+ \in \mathbb{R}$ such that $\mathbf{P}\left(\ell^* \leq X_i < \ell_\epsilon^+\right) = \epsilon$. Show that $\mathbf{P}\left(\ell_\epsilon^+ \leq X_{\min}^n\right) = \left(1-\epsilon\right)^n$.
2. Show that if $n > \frac{1}{\epsilon} \log\left(\frac{1}{\delta}\right)$, then $\mathbf{P}\left(\ell_\epsilon^+ \leq X_{\min}^n\right) < \delta$. *Hint:* $1 - z \leq \exp\left(-z\right)$.
3. Show that $\mathcal{C}$ is efficiently PAC-learnable from $\mathcal{C}$. *Hint: Prove* $\mathbf{P}\left(\mathcal{R}\left(\hat{c}_n\right) > \epsilon\right) = \mathbf{P}\left(\ell_\epsilon^+ \leq X_{\min}^n\right)$.