Assignment 1

Deadline: Wednesday, 22^{nd} of May 2024

Upload your solutions at: https://tinyurl.com/AML-2024-ASSIGNMENT1

1. **(0.5 points)** Give examples for:

a) a finite hypothesis class $\mathcal{H}$ with VCdim$(\mathcal{H}) = 2024$. Justify your choice. **(0.25 points)**

b) an infinite hypothesis class $\mathcal{H}$ with VCdim$(\mathcal{H}) = 2024$. Justify your choice. **(0.25 points)**

2. **(1 point)** Let $\mathcal{X} = \mathbb{R}^2$ and consider $\mathcal{H}_\alpha$ the set of concepts defined by the area inside a right triangle ABC with two catheti AB and AC parallel to the axes (Ox and Oy), and with the ratio AB/AC $= \alpha$ (fixed constant > 0). Consider the realizability assumption. Show that the class $\mathcal{H}_\alpha$ is (ϵ, δ)-PAC learnable by giving an algorithm A and determining an upper bound on the sample complexity $m_H(\epsilon, \delta)$ such that the definition of PAC-learnability is satisfied.

3. **(1 point)** Consider $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3$, where:

$$\mathcal{H}_1 = \{h_a : \mathbb{R} \to \{0,1\} \mid h_a(x) = \mathbf{1}_{[x \geq a]}(x) = \mathbf{1}_{[a,+\infty)}(x), a \in \mathbb{R}\},$$

$$\mathcal{H}_2 = \{h_b : \mathbb{R} \to \{0,1\} \mid h_b(x) = \mathbf{1}_{[x < b]}(x) = \mathbf{1}_{(-\infty,b)}(x), b \in \mathbb{R}\},$$

$$\mathcal{H}_3 = \{h_{c,d} : \mathbb{R} \to \{0,1\} \mid h_{c,d}(x) = \mathbf{1}_{[c \leq x \leq d]}(x) = \mathbf{1}_{[c,d]}(x), c, d \in \mathbb{R}\}.$$

Consider the realizability assumption. Compute VCdim$(\mathcal{H})$.

4. **(1 point)** Consider $\mathcal{H}$ the class of 3-piece classifiers (signed intervals):

$$\mathcal{H} = \{h_{a,b,s} : \mathbb{R} \to \{-1,1\} \mid a \leq b, s \in \{-1,1\}\}, \ where \ h_{a,b,s}(x) = \begin{cases} s, & x \in [a,b] \\ -s, & x \notin [a,b] \end{cases}$$

a. Compute the shattering coefficient $\tau_H(m)$ of the growth function for $m \geq 0$ for hypothesis class $\mathcal{H}$. **(0.75 points)**

b. Compare your result with the general upper bound for the growth functions and show that $\tau_H(m)$ obtained at previous point a is not equal with the upper bound. **(0.1 points)**

c. Does there exist a hypothesis class $\mathcal{H}$ for which the shattering coefficient $\tau_H(m)$ is equal to the general upper bound (over or another domain $\mathcal{X}$)? If your answer is yes please provide an example, if your answer is no please provide a justification. **(0.15 points)**

5. **(1 point)** Let $\mathcal{H} = \{h : \mathbb{R} \to \{0,1\} \; | h_\theta(x) = \mathbf{1}_{[\theta,\theta+1]\cup[\theta+2,\infty)}(x), \theta \in \mathbb{R}\}$.
Compute VCdim($\mathcal{H}$).

6. **(1 point)** A decision list may be thought of as an ordered sequence of if-then-else statements. The sequence of conditions in the decision list is tested in order, and the answer associated with the first satisfied condition is output.

More formally, a *k-decision list* over the boolean variables $x_1, x_2, \ldots, x_n$ is an ordered sequence $L = \{(c_1, b_1), (c_2, b_2), \ldots, (c_l, b_l)\}$ and a bit b, in which each c_i is a conjunction of at most k literals over $x_1, x_2, \ldots, x_n$ and each $b_i \in \{0,1\}$. For any input $a \in \{0,1\}^n$, the value $L(a)$ is defined to be b_j where j is the smallest index satisfying $c_j(a) = 1$; if no such index exists, then $L(a) = b$. Thus, b is the "default" value in case a falls off the end of the list. We call b_i the bit associated with the condition c_i.

The next figure shows an example of a *2-decision list* along with its evaluation on a particular input.

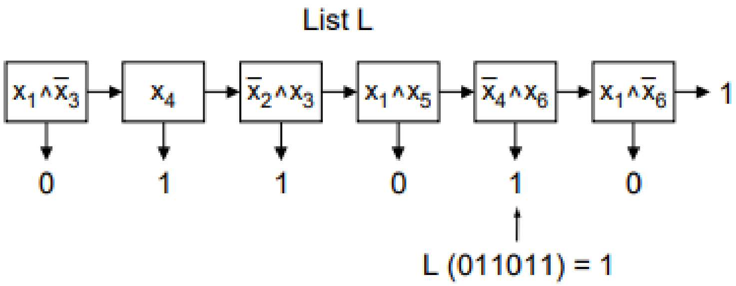

Figure 1: *A 2-decision list and the path followed by an input. Evaluation starts at the leftmost item and continues to the right until the first condition is satisfied, at which point the binary value below becomes the final result of the evaluation.*

Show that the VC dimension of 1-decision lists over $\{0,1\}^n$ is lower and upper bounded by linear functions, by showing that there exists $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ such that:

$$\alpha \cdot n + \beta \leq VCdim(\mathcal{H}_{1-decision\ list}) \leq \gamma \cdot n + \delta$$

Hint: Show that 1-decision lists over $\{0,1\}^n$ compute linearly separable functions (halfspaces).

Ex-officio: 0.5 points