Assignment 2

Deadline: Wednesday, 14^{th} of June, 23:59

Upload your solutions as a zip archive at:
https://tinyurl.com/AML-2023-ASSIGNMENT2

1. **(1.25 points)** Consider $\mathcal{H}$ the following hypothesis class :

$$\mathcal{H} = \{h_a : \mathbb{R} \to \{0,1\} \mid a > 0, a \in \mathbb{R},\ where\ h_a(x) = \mathbf{1}_{[-a,a]}(x) = \begin{cases} 1, & x \in [-a,a] \\ 0, & x \notin [-a,a] \end{cases} \}$$

a. Compute the shattering coefficient $\tau_H(m)$ of the growth function for $m \geq 0$ for hypothesis class $\mathcal{H}$. **(1 point)**

b. Compare your result from the previous point with the general upper bound given by the Sauer lemma. Are they equal or different? **(0.25 points)**

2. **(1.5 points)** Consider de concept class C_a formed by the union of two closed intervals $[a, a+1] \cup [a+2, a+4]$, where $a \in \mathbb{R}$. Give an efficient ERM algorithm for learning the concept class C_a and compute its complexity for each of the following cases:

a. realizable case. **(1 point)**

b. agnostic case. **(0.5 point)**

3. **(1.25 points)** Consider a modified version of the AdaBoost algorithm that runs for exactly three rounds as follows:

- the first two rounds run exactly as in AdaBoost (at round 1 we obtain distribution $\mathbf{D}^{(1)}$, weak classifier h_1 with error ϵ_1; at round 2 we obtain distribution $\mathbf{D}^{(2)}$, weak classifier h_2 with error ϵ_2).
- in the third round we compute for each $i = 1, 2, \ldots, m$:

$$\mathbf{D}^{(3)}(i) = \begin{cases} \frac{D^{(1)}(i)}{Z}, & if\ h_1(x_i) \neq h_2(x_i) \\ 0, & otherwise \end{cases}$$

  where Z is a normalization factor such that $\mathbf{D}^{(3)}$ is a probability distribution.
- obtain weak classifier h_3 with error ϵ_3.
- output the final classifier $h_{final}(x) = sign(h_1(x) + h_2(x) + h_3(x))$.

Assume that at each round $t = 1, 2, 3$ the weak learner returns a weak classifier h_t for which the error ϵ_t satisfies $\epsilon_t \leq \frac{1}{2} - \gamma_t, \gamma_t > 0$.

a. What is the probability that the classifier h_1 (selected at round 1) will be selected again at round 2? Justify your answer. **(0.25 points)**

b. Consider $\gamma = min\{\gamma_1, \gamma_2, \gamma_3\}$. Show that the training error of the final classifier h_{final} is at most $\frac{1}{2} - \frac{3}{2}\gamma + 2\gamma^3$ and show that this is strictly smaller than $\frac{1}{2} - \gamma$. **(1 point)**

4. **(1 point)** Consider H^d_{2DNF} the class of 2-term disjunctive normal form formulae consisting of hypothesis of the form $h : \{0,1\}^d \to \{0,1\}$,

$$h(x) = A_1(x) \vee A_2(x)$$

where $A_i(x)$ is a Boolean conjunction of literals H^d_{conj}.

It is known that the class H^d_{2DNF} is not efficient properly learnable but can be learned improperly considering the class H^d_{2CNF}. Give a γ-weak-learner algorithm for learning the class H^d_{2DNF} which is not a strong PAC learning algorithm for H^d_{2DNF} (like the one considering H^d_{2CNF}). Prove that this algorithm is a γ-weak-learner algorithm for H^d_{2DNF}.

Hint: Find an algorithm that returns $h(x) = 0$ or the disjunction of 2 literals.

Ex-oficio: 0.5 points.