Machine Learning, Spring 2023
Homework 4

Due on 23:59 April 27, 2023

1 Understanding VC dimension (50 points)

In this part, you need to complete some mathematical proofs about VC dimension. Suppose the hypothesis set

$$\mathcal{H} = \{f(x, \alpha) = \text{sign }(\sin(\alpha x))|, \alpha \in \Re\}$$

where x and f are feature and label, respectively.

- Show that $\mathcal{H}$ cannot shatter the points $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4$. (20 points)

  (Key: Mathematically, you need to show that there exists y_1, y_2, y_3, y_4, for any $\alpha \in \Re$, $f(x_i) \neq y_i, i = 1, 2, 3, 4$, for example, $+1, +1, -1, +1$)

- Show that the VC dimension of $\mathcal{H}$ is ∞. (Note the difference between it and the first question) (30 points)

  (Key: Mathematically, you have to prove that for any label sets $y_1, \cdots, y_m, m \in \mathbb{N}$, there exists $\alpha \in \Re$ and $x_i, i = 1, 2, \cdots, m$ such that $f(x; \alpha)$ can generate this set of labels. Consider the points $x_i = 10^{-i}$...)

2 Bias-variance decomposition (50 points)

When there is noise in the data, $E_{out}(g^{(\mathcal{D})}) = \mathbb{E}_{\mathbf{x},y}\left[(g^{(\mathcal{D})}(\mathbf{x}) - y(\mathbf{x}))^2\right]$, where $y(\mathbf{x}) = f(\mathbf{x}) + \epsilon$. If ϵ is a zero-mean noise random variable with variance σ^2, show that the bias-variance decomposition becomes

$$\mathbb{E}_{\mathcal{D}}\left[E_{out}(g^{(\mathcal{D})})\right] = \sigma^2 + \text{bias} + \text{var}$$