Machine Learning, 2024 Spring
Assignment 5

Name: Zhou Shouchen
Student ID: 2021533042

Notice

Plagiarizer will get 0 points.
LaTeXis highly recommended. Otherwise you should write as legibly as possible.

Problem 1

Which of the following are possible growth functions $m_{\mathcal{H}}(N)$ for some hypothesis set:

$$1+N; 1+N+\frac{N(N-1)}{2}; 2^N; 2^{\lfloor\sqrt{N}\rfloor}; 2^{\lfloor\frac{N}{2}\rfloor}; 1+N+\frac{N(N-1)(N-2)}{6}$$

Solution

1. $1+N$:

2. $1+N+\dfrac{N(N-1)}{2}$:

3. 2^N:

4. $2^{\lfloor\sqrt{N}\rfloor}$:

5. $2^{\lfloor\frac{N}{2}\rfloor}$:

6. $1+N+\dfrac{N(N-1)(N-2)}{6}$:

Problem 2

For an $\mathcal{H}$ with $d_{\mathrm{vc}} = 10$, what sample size do you need (as prescribed by the generalization bound) to have a 95% confidence that your generalization error is at most 0.05 ?

Solution

Problem 3

Let $\mathcal{H} = \{h_1, h_2, \dots, h_M\}$ with some finite M. Prove that $d_{\text{vc}}(\mathcal{H}) \leq \log_2 M$.

Solution

Problem 4

Let $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_K$ be K hypothesis sets with finite VC dimension d_{vc} . Let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2 \cup \ldots \cup \mathcal{H}_K$ be the union of these models. Show that $d_{\text{vc}}(\mathcal{H}) < K\,(d_{\text{vc}} + 1)$.

Solution

Problem 5

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

$$\mathcal{H} = \{f(x,\alpha) = \text{sign}(\sin(\alpha x)) \mid, \alpha \in \mathbb{R}\}$$

where x and f are feature and label, respectively.

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

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

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

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

Solution

1.

2.