

Learning theory

Machine Learning II
2021-2022 - UMONS
Souhaib Ben Taieb

1 Exercise 1

Solve Problem 2.8 in LFD..

Problem 2.8 Which of the following are possible growth functions $m_{\mathcal{H}}(N)$ for some hypothesis set:

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

Figure 1: Source: Abu-Mostafa et al. Learning from data. AMLbook.

2 Exercise 2

Solve Problem 2.3 in LFD.

Problem 2.3 Compute the maximum number of dichotomies, $m_{\mathcal{H}}(N)$, for these learning models, and consequently compute d_{VC} , the VC dimension.

- (a) Positive or negative ray: \mathcal{H} contains the functions which are $+1$ on $[a, \infty)$ (for some a) together with those that are $+1$ on $(-\infty, a]$ (for some a).
- (b) Positive or negative interval: \mathcal{H} contains the functions which are $+1$ on an interval $[a, b]$ and -1 elsewhere or -1 on an interval $[a, b]$ and $+1$ elsewhere.
- (c) Two concentric spheres in \mathbb{R}^d : \mathcal{H} contains the functions which are $+1$ for $a \leq \sqrt{x_1^2 + \dots + x_d^2} \leq b$.

Figure 2: Source: Abu-Mostafa et al. Learning from data. AMLbook.

3 Exercise 3

Solve Problem 2.4 in LFD.

Problem 2.4 Show that $B(N, k) = \sum_{i=0}^{k-1} \binom{N}{i}$ by showing the other direction to Lemma 2.3, namely that

$$B(N, k) \geq \sum_{i=0}^{k-1} \binom{N}{i} .$$

To do so, construct a specific set of $\sum_{i=0}^{k-1} \binom{N}{i}$ dichotomies that does not shatter any subset of k variables. *[Hint: Try limiting the number of -1 's in each dichotomy.]*

Figure 3: Source: Abu-Mostafa et al. Learning from data. AMLbook.

4 Exercise 4

Solve Problem 2.5 in LFD.

Problem 2.5 Prove by induction that $\sum_{i=0}^D \binom{N}{i} \leq N^D + 1$, hence

$$m_{\mathcal{H}}(N) \leq N^{d_{\text{vc}}} + 1.$$

Figure 4: Source: Abu-Mostafa et al. Learning from data. AMLbook.

5 Exercise 5

Solve Problem 2.13 in LFD..

Problem 2.13

- (a) Let $\mathcal{H} = \{h_1, h_2, \dots, h_M\}$ with some finite M . Prove that $d_{\text{VC}}(\mathcal{H}) \leq \log_2 M$.
- (b) For hypothesis sets $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_K$ with finite VC dimensions $d_{\text{VC}}(\mathcal{H}_k)$, derive and prove the tightest upper and lower bound that you can get on $d_{\text{VC}}(\cap_{k=1}^K \mathcal{H}_k)$.
- (c) For hypothesis sets $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_K$ with finite VC dimensions $d_{\text{VC}}(\mathcal{H}_k)$, derive and prove the tightest upper and lower bounds that you can get on $d_{\text{VC}}(\cup_{k=1}^K \mathcal{H}_k)$.

Figure 5: Source: Abu-Mostafa et al. Learning from data. AMLbook.