Reference Solution to the Quiz 7

Xiangyu Yang

April 15, 2020

1 Lecture 13

According to Theorem 7.1 shown in the course slide, please derive the following sample complexity for the consistent learner, which reads

$$m \geq \frac{1}{\epsilon}\left[\ln(|H|) + \ln(\frac{1}{\delta})\right]. \tag{1}$$

Proof. By Theorem 7.1, and let $\delta > 0$ be an upper bound on the probability of not exhausting the version space, so

$$\Pr(\exists h \in VS_{H,D}, err_D(h) \geq \epsilon) \leq |H|e^{-\epsilon m} \leq \delta. \tag{2}$$

Focus on the second inequality of (2), we have

$$|H|e^{-\epsilon m} \leq \delta \Longleftrightarrow \ln |H|e^{-\epsilon m} \leq \ln \delta. \tag{3}$$

Hence, after some simple algebraic manipulations, we can easily obtain the desired inequality (1). This completes the proof. □

2 Lecture 14

Below is the definition of $H[S]$ and $H[m]$, Correspondingly:

$H[S] = \{(h(x_1), \cdots, h(x_m)) \mid h \in \mathcal{H}\}$,

$H(m) = \max\limits_{\|S\|=m} \{ |H(S)| \mid \forall S \in \mathcal{X} \}$.

Solution:

1. $H(S) = \{\overbrace{(+,+,-,-), (+,-,-,-), \cdots, (-,-,-,-)}^{16 \text{ Combinations}}\}$

Also, you can use label +1/-1 or +1/0 besides +/-.

2. $H(4) = \max\limits_{\|S\|=4} \{ |H(S)| \mid \forall S \in \mathcal{X} \} = 2^4 = 16$

3. VC dim$(\mathcal{H}) \geq 4$ is True.

Because it suffices to show that 5 points can be shattered. For example,