Chapter 1 Quiz Practice Problems

Problem 1. For each statement below, circle `True` if the statement is known to be true, `False` if the statement is known to be false, and `Open` if the statement is not known to be either true or false. Ensure that you pay careful attention to the formal definitions of asymptotic notation in your responses.

1. `True` `False` `Open` Let $\mathcal{H}$ be the perceptron hypothesis class. Assume there exists some $h \in \mathcal{H}$ with $E_{\text{out}}(h) = 0$. Then after a finite number of iterations, the PLA is guaranteed to output a hypothesis g satisfying $E_{\text{in}}(g) = 0$.

2. `True` `False` `Open` Let $\mathcal{H}$ be the perceptron hypothesis class. Assume there exists some $h \in \mathcal{H}$ with $E_{\text{in}}(h) = 0$. Then after a finite number of iterations, the PLA is guaranteed to output a hypothesis g satisfying $E_{\text{out}}(g) = 0$.

3. `True` `False` `Open` Let $\mathcal{H}$ be the perceptron hypothesis class. Assume there exists some $h \in \mathcal{H}$ with $E_{\text{in}}(h) = 0$. Then the PLA will terminate.

4. `True` `False` `Open` Let $\mathcal{H}$ be a hypothesis class. If there exists a hypothesis $h \in \mathcal{H}$ such that $E_{\text{out}}(h) = 0$, then $\mathcal{H}$ must be finite.

5. `True` `False` `Open` If your dataset is linearly separable, then the PLA is guaranteed to terminate.

6. `True` `False` `Open` Let
$$\mathcal{H}_{\text{axis2}} = \Big\{\mathbf{x} \mapsto \sigma \operatorname{sign}(x_i) : \sigma \in \{+1, -1\}, i \in [d]\Big\},$$
and
$$\mathcal{H}_{\text{L2-4}} = \Big\{\mathbf{x} \mapsto [\![\|\mathbf{x}\|_2 \geq \alpha]\!] : \alpha \in \{1, 2, 3, 4\}\Big\}.$$
Let $g_{\text{axis2}} \in \mathcal{H}_{\text{axis2}}$ and $g_{\text{L2-4}} \in \mathcal{H}_{\text{L2-4}}$ be the outputs of the TEA algorithm on their respective hypothesis classes. Then for $d > 100$, $\mathcal{H}_{\text{axis2}} \subset \mathcal{H}_{\text{L2-4}}$.

7. `True` `False` `Open` Let
$$\mathcal{H}_{\text{axis2}} = \Big\{\mathbf{x} \mapsto \sigma \operatorname{sign}(x_i) : \sigma \in \{+1, -1\}, i \in [d]\Big\},$$
and
$$\mathcal{H}_{\text{L2-4}} = \Big\{\mathbf{x} \mapsto [\![\|\mathbf{x}\|_2 \geq \alpha]\!] : \alpha \in \{1, 2, 3, 4\}\Big\}.$$
Let $g_{\text{axis2}} \in \mathcal{H}_{\text{axis2}}$ and $g_{\text{L2-4}} \in \mathcal{H}_{\text{L2-4}}$ be the outputs of the TEA algorithm on their respective hypothesis classes. Then for $d > 100$, the following inequality is guaranteed to hold: $E_{\text{in}}(g_{\text{axis2}}) \leq E_{\text{in}}(g_{\text{L2-4}})$.

8. `True` `False` `Open` Let

$$\mathcal{H}_{\text{axis2}} = \left\{ \mathbf{x} \mapsto \sigma \, \text{sign}(x_i) : \sigma \in \{+1, -1\}, i \in [d] \right\},$$

and

$$\mathcal{H}_{\text{L2-4}} = \left\{ \mathbf{x} \mapsto [\![\|\mathbf{x}\|_2 \geq \alpha]\!] : \alpha \in \{1, 2, 3, 4\} \right\}.$$

Let $g_{\text{axis2}} \in \mathcal{H}_{\text{axis2}}$ and $g_{\text{L2-4}} \in \mathcal{H}_{\text{L2-4}}$ be the outputs of the TEA algorithm on their respective hypothesis classes. Then for learning problems with large d, the finite hypothesis class generalization theorem predicts that $g_{\text{L2-4}}$ will have better generalization accuracy than g_{axis2} with high probability.

9. `True` `False` `Open` Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hypothesis classes satisfying $\mathcal{H}_1 \subset \mathcal{H}_2$. Let g_1 be the result of the TEA algorithm on $\mathcal{H}_1$ and g_2 be the result of TEA on $\mathcal{H}_2$. Then the following inequality is guaranteed to hold: $E_{\text{in}}(g_1) \leq E_{\text{in}}(g_2)$.

10. `True` `False` `Open` Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hypothesis classes satisfying $\mathcal{H}_1 \subset \mathcal{H}_2$. Let g_1 be the result of the TEA algorithm on $\mathcal{H}_1$ and g_2 be the result of TEA on $\mathcal{H}_2$. Then Hoeffding's inequality predicts that with high probability, $E_{\text{test}}(g_1) \leq E_{\text{test}}(g_2)$.

11. `True` `False` `Open` Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hypothesis classes satisfying $\mathcal{H}_1 \subset \mathcal{H}_2$. Let g_1 be the result of the TEA algorithm on $\mathcal{H}_1$ and g_2 be the result of TEA on $\mathcal{H}_2$. Then the finite hypothesis class generalization theorem predicts that with high probability, $E_{\text{out}}(g_1) \leq E_{\text{out}}(g_2)$.

12. `True` `False` `Open` Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hypothesis classes satisfying $\mathcal{H}_1 \subset \mathcal{H}_2$. Let g_1 be the result of the TEA algorithm on $\mathcal{H}_1$ and g_2 be the result of TEA on $\mathcal{H}_2$. Then the finite hypothesis class generalization theorem predicts that with high probability, $E_{\text{out}}(g_1) \leq E_{\text{out}}(g_2)$.

13. `True` `False` `Open` For all $d > 100$, we have that $|\mathcal{H}_{\text{binary}}| < |\mathcal{H}_{\text{axis}}|$.

14. `True` `False` `Open` For all $d > 100$, we have that $\mathcal{H}_{\text{binary}} \subseteq \mathcal{H}_{\text{axis}}$.

15. `True` `False` `Open` Let g_{axis} be the result of running TEA on $\mathcal{H}_{\text{axis}}$, g_{binary} be the result of running TEA on $\mathcal{H}_{\text{binary}}$, g_{union} be the result of running TEA on $\mathcal{H}_{\text{union}} = \mathcal{H}_{\text{axis}} \cup \mathcal{H}_{\text{binary}}$. Then we have that for all datasets,

$$E_{\text{in}}(g_{\text{axis}}) \leq E_{\text{in}}(g_{\text{binary}}) \leq E_{\text{in}}(g_{\text{union}}).$$