Chapter 3 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.

1. `True` `False` `Open` You have trained a perceptron model that has high generalization error. VC theory predicts that applying the PCA feature embedding with a low output dimension will reduce the generalization error with high probability.

2. `True` `False` `Open` You have trained a perceptron model that has high in sample error. VC theory predicts that applying the PCA feature embedding with a low output dimension will reduce the in sample error with high probability.

3. `True` `False` `Open` Let $\mathcal{H}_{\text{perceptron}}$ be the perceptron hypothesis class and let $\mathcal{H}_{\text{stump}}$ be the decision stump hypothesis class. VC theory predicts that the generalization error of $\mathcal{H}_{\text{stump}}$ will be lower than the generalization error of $\mathcal{H}_{\text{perceptron}}$ with high probability.

4. `True` `False` `Open` Let $\mathcal{H}_{\tilde{d}}$ be the perceptron hypothesis class with the PCA feature map of dimension $\tilde{d}$. If k is a breakpoint for $\mathcal{H}_1$, then k will also be a breakpoint for $\mathcal{H}_2$.

5. `True` `False` `Open` Let $\mathcal{H}_Q$ be the perceptron hypothesis class with the polynomial feature map of degree Q. If k is a breakpoint for $\mathcal{H}_1$, then k will also be a breakpoint for $\mathcal{H}_2$.

6. `True` `False` `Open` VC theory predicts that when using the perceptron hypothesis class, centering your data points results in a better generalization error with high probability.

7. `True` `False` `Open` Let $\mathcal{H}_{\text{perceptron}}$ be the perceptron hypothesis class and let $\mathcal{H}_{\text{stump}}$ be the decision stump hypothesis class. Let $g_{\text{perceptron}} \in \mathcal{H}_{\text{perceptron}}$ and $g_{\text{stump}} \in \mathcal{H}_{\text{stump}}$ be the empirical risk minimizers. VC theory predicts that $E_{\text{out}}(g_{\text{stump}}) \leq E_{\text{out}}(g_{\text{perceptron}})$ with high probability.

8. `True` `False` `Open` Let $\mathcal{H}$ be the perceptron hypothesis class and let $\mathcal{H}_\Phi$ be the perceptron hypothesis class with the polynomial feature map of degree 2. Let $g \in \mathcal{H}$ and $g_\Phi \in \mathcal{H}_\Phi$ be the empirical risk minimizers trained on a very large dataset. Then we are guaranteed that $E_{\text{in}}(g) \geq E_{\text{in}}(g_\Phi)$.

9. True False Open Let $g_{\text{axis2}} \in \mathcal{H}_{\text{axis2}}$ be the empirical risk minimizer for the $\mathcal{H}_{\text{axis2}}$ hypothesis class and $g_{\text{stump}} \in \mathcal{H}_{\text{stump}}$ be the empirical risk minimizer for the decision stump hypothesis class. Then we are guaranteed that $E_{\text{in}}(g_{\text{axis2}}) \leq E_{\text{in}}(g_{\text{stump}})$.

10. True False Open Let X be a dataset shattered by the perceptron hypothesis class. Then X is guaranteed to also be shattered by the perceptron hypothesis class with the polynomial feature map of degree 7.

11. True False Open Let X be a dataset shattered by the perceptron hypothesis class. Then X is guaranteed to also be shattered by the decision stump hypothesis class.

12. True False Open Let X be a dataset shattered by the decision stump hypothesis class. Then X is guaranteed to also be shattered by the perceptron hypothesis class with polynomial feature map of degree 2.

13. True False Open There exists a dataset of size $N = 2$ in $d = 2$ dimensions that can be shattered by the perceptron hypothesis class but cannot be shattered by the $\mathcal{H}_{\text{axis2}}$ hypothesis class.

14. True False Open There exists a dataset of size $N = 2$ in $d = 2$ dimensions that can be shattered by the perceptron hypothesis class but cannot be shattered by the decision stump hypothesis class.

15. True False Open Define the hypothesis class of concentric circles with the feature map Φ as

$$\mathcal{H}_{\text{circles},\Phi} = \left\{ \mathbf{x} \mapsto [\![\|\Phi(\mathbf{x})\|_2 \geq \alpha]\!] : \alpha \in \mathbb{R} \right\}.$$

You have trained an empirical risk minimizer on this hypothesis class using the polynomial feature map of degree $Q = 20$; the resulting hypothesis has high generalization error. VC theory predicts that reducing the degree of the polynomial to $Q = 2$ will also reduce the generalization error with high probability.

16. True False Open Define the $\mathcal{H}_{\text{axis2}}$ hypothesis class with the feature map $\Phi : d \to \tilde{d}$ as

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

You have trained an empirical risk minimizer on this hypothesis class using the polynomial feature map of degree $Q = 20$; the resulting hypothesis has high generalization error. VC theory predicts that reducing the degree of the polynomial to $Q = 2$ will also reduce the generalization error with high probability.