Quiz: Chapter 3 (Section 3.4)

Printed Name:

Quiz rules:

1. You MAY use:
   (a) any notes (handwritten, printed, or electronic),
   (b) any computer programs (including websites like WolframAlpha or ChatGPT), and
   (c) any additional scratch paper.
2. You MAY NOT communicate with another student.
3. If you finish the quiz early, stay seated. I will collect all the quizzes at the same time.

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. You will receive +2 point for each correct answer, **-1 points for each incorrect answer**, and 0 points for each blank answer.

1. `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. If k is a breakpoint for $\mathcal{H}_{\text{stump}}$, then k is guaranteed to be a breakpoint for $\mathcal{H}_{\text{perceptron}}$ as well.

2. `True` `False` `Open` Let $\mathcal{H}_{\Phi_2}$ be the perceptron hypothesis class with the 2nd degree polynomial feature embedding, and $\mathcal{H}_{\Phi_4}$ be perceptron hypothesis class with the 4th degree polynomial feature embedding. Let $g_{\Phi_2} \in \mathcal{H}_{\Phi_2}$ and $g_{\Phi_4} \in \mathcal{H}_{\Phi_4}$ be the empirical risk minimizers. It is guaranteed that $E_{\text{in}}(g_{\Phi_2}) \leq E_{\text{in}}(g_{\Phi_4})$.

3. `True` `False` `Open` Let $\mathcal{H}_{\Phi_2}$ be the perceptron hypothesis class with the PCA feature embedding with output dimension 2, and $\mathcal{H}_{\Phi_4}$ be perceptron hypothesis class with the PCA feature embedding with output dimension 4. Let $g_{\Phi_2} \in \mathcal{H}_{\Phi_2}$ and $g_{\Phi_4} \in \mathcal{H}_{\Phi_4}$ be the empirical risk minimizers. VC theory predicts that with high probability, the generalization error of g_{Φ_2} will be less than the generalization error of g_{Φ_4}.

4. `True` `False` `Open` VC theory predicts that when using the $\mathcal{H}_{\text{axis2}}$ hypothesis class, centering your data points results in a better generalization error with high probability.