

More Midterm 3 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 | If your model is overfitting, then VC theory predicts that you should increase the size of your VC dimension. |
| 2. True | False | Open | If your model is underfitting, VC theory predicts that you should increase the size of your VC dimension. |
| 3. True | False | Open | If you are training a logistic regression model that has low in-sample error E_{in} but high generalization error $ E_{\text{in}} - E_{\text{out}} $, then VC theory predicts that increasing the number of data points N in the training set will result in better performance than increasing the number of dimensions d . |
| 4. True | False | Open | If you are training a logistic regression model with the polynomial kernel that has high in-sample error E_{in} and low generalization error $ E_{\text{in}} - E_{\text{out}} $, then VC theory predicts that increasing the degree of the kernel is more likely to improve performance than decreasing the degree of the kernel. |
| 5. True | False | Open | You have trained a logistic regression model with L2 regularization. If your training set size N increases, then the optimal soft order constraint regularization hyperparameter C will also increase. |
| 6. True | False | Open | You have trained a logistic regression model with L1 regularization. If your training set size N increases, then the optimal augmented error regularization hyperparameter λ will also increase. |
| 7. True | False | Open | You have trained a logistic regression model with L1 regularization. If your number of feature dimensions d increases, then the optimal soft order constraint regularization hyperparameter C will also increase. |
| 8. True | False | Open | You have trained a logistic regression model with elastic net regularization. If your number of feature dimensions d increases, then the optimal augmented error regularization hyperparameter λ will also increase. |
| 9. True | False | Open | When training a logistic regression model, if you want the weight vector to be sparse, then you should prefer L1 regularization to L2 regularization. |

10. True	False	Open	When training a logistic regression model using L2 regularization, increasing the value of the augmented error regularization hyperparameter λ increases the VC dimension.
11. True	False	Open	The elastic net regularization function cannot be applied with the polynomial kernel to a logistic regression model.
12. True	False	Open	You have trained a logistic regression model with the PCA kernel and used a validation set to determine that the optimal output dimension k is 200. If your training set size N increases, then VC theory predicts that the optimal value for k will increase.
13. True	False	Open	You have trained a MLP with the ReLU activation function and used a validation set to determine that the optimal number of layers is 5 and optimal width is 100. If your training set size N increases, and you keep the number of layers the same, then VC theory predicts that the optimal width of those layers will increase.
14. True	False	Open	You have trained a boosted decision stump model and used a validation set to determine that the optimal number of base classifiers T is 1000. If instead of training a decision stump, you train a decision tree of depth 2, then the optimal number of base classifiers will increase.
15. True	False	Open	The VC dimension of neural networks with the ReLU activation function is $\Theta(Ek \log(E))$, where k and E is as-defined in the notes.
16. True	False	Open	The VC dimension of neural networks with the identity activation function is $O(dE)$.
17. True	False	Open	Assume you are training a boosted decision stump model on a dataset with $N = 10^6$ and $d = 10^6$. Then in the limit as the number of base models T approaches infinity, the training error is guaranteed to approach 0 for all possible datasets.
18. True	False	Open	Assume you are training an SVM with the random features kernel with output dimension d' . Then in the limit as d' approaches infinity, the training error is guaranteed to approach 0 for all possible datasets.
19. True	False	Open	In vowpal wabbit, increasing the <code>--bit_precision</code> hyperparameter increases the model's VC dimension.

20.	True	False	Open	In vowpal wabbit, increasing the <code>--hash_seed</code> hyperparameter increases the model's VC dimension.
21.	True	False	Open	In vowpal wabbit, increasing the <code>--ngrams</code> hyperparameter increases the model's VC dimension.
22.	True	False	Open	In scikit-learn's <code>sklearn.tree.DecisionTreeClassifier</code> model, increasing the value of the hyperparameter <code>max_depth</code> increases the VC dimension.
23.	True	False	Open	You have trained a scikit-learn <code>sklearn.tree.DecisionTreeClassifier</code> model, but it is underfitting. VC theory predicts that increasing the value of <code>min_samples_split</code> will be more likely to improve performance than decreasing this value.
24.	True	False	Open	You are training a scikit-learn <code>sklearn.ensemble.AdaBoostClassifier</code> model with <code>base_estimator</code> set to <code>sklearn.tree.DecisionTreeClassifier</code> . In order to keep the VC dimension of your model constant, if you increase the value of <code>n_estimators</code> for the <code>AdaBoostClassifier</code> , then you should also increase the value of <code>max_depth</code> for the <code>DecisionTreeClassifier</code> .
25.	True	False	Open	You are training a scikit-learn <code>sklearn.ensemble.AdaBoostClassifier</code> model with <code>base_estimator</code> set to <code>sklearn.tree.DecisionTreeClassifier</code> . In order to keep the VC dimension of your model constant, if you increase the value of <code>min_samples_split</code> for the <code>DecisionTreeClassifier</code> then you should also increase the value of <code>n_estimators</code> for the <code>AdaBoostClassifier</code> .
26.	True	False	Open	You are using transfer learning to train the final layer of a deep neural network for your specific task. The <code>ResNet50</code> model has 50 hidden layers and the <code>ResNet18</code> model has only 18 hidden layers, and in both cases all layers have the same width. VC theory predicts that if you use the <code>ResNet50</code> model you will have a higher generalization error than if you use the <code>ResNet18</code> model.
27.	True	False	Open	You are using transfer learning to train the final layer of a deep neural network for your specific task. The <code>VGG13</code> model has 13 hidden layers and the <code>VGG19</code> model has 19 hidden layers, and in both cases all layers have the same width. VC theory predicts that if you use the <code>VGG13</code> model you will have a higher generalization error than if you use the <code>VGG19</code> model.