Quiz 4

Time limit. 30 minutes.

Collaboration and materials. You must complete this quiz individually. You may not use any materials (physical or electronic) besides both sides of one sheet of 8.5x11-inch paper with 1-inch margins and the equivalent of 10-point font.

Questions. This quiz has ten multiple-choice questions. Some questions require you to select exactly one of the answer choices, while others require you to select all of the answer choices that apply. Questions of the latter kind always end with "Select all that apply."

Scoring. Each question is weighted equally. For questions requiring you to select one of the answer choices, no partial credit will be awarded. For questions requiring you to select all of the answer choices that apply, partial credit will be awarded for each correct answer selected while no points will be awarded if no correct answers are chosen or if any incorrect answers are selected.

Submission. You will receive a bubble sheet for your answers. Please print your full name as it appears on Gradescope (please no cursive), your student ID, and today's date (November 9). You may leave the "Section" box blank. **Your version is A. Please check that this matches the pre-bubbled version number at the top of the bubble sheet.** For each question, please fill in the appropriate bubbles completely using either pencil or blue/black pen. If you have filled in a bubble with pen but have changed your mind, you can cross out that bubble with an X. Note that the answer choices are presented in the order A, B, C, D, E.

Typesetting math: 100%

1

1 point

You are carrying out bagging on a training dataset with just three observations. When you draw one bootstrap sample, what is the probability that observation 1 is out of bag? [Hint: Start by considering drawing one bootstrap observation at a time, and find the probability at each step that observation 1 is not chosen. Then, use the fact that the probability of a sequence of independent events is the product of the probabilities of the events.]

○ 2/9

○ 8/27

○ 1/27

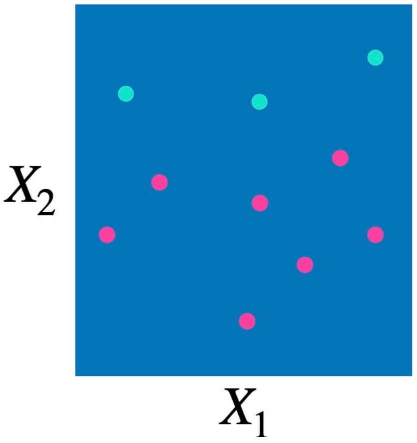
○ 10/27

○ 19/27

2

1 point

Below is a dataset on which you would like to train a classification tree:



Let D be the decrease in the total Gini index caused by the first split in the classification tree. To which of the following intervals does D belong?

○ [0, 0.2)

○ [0.4, 0.6)

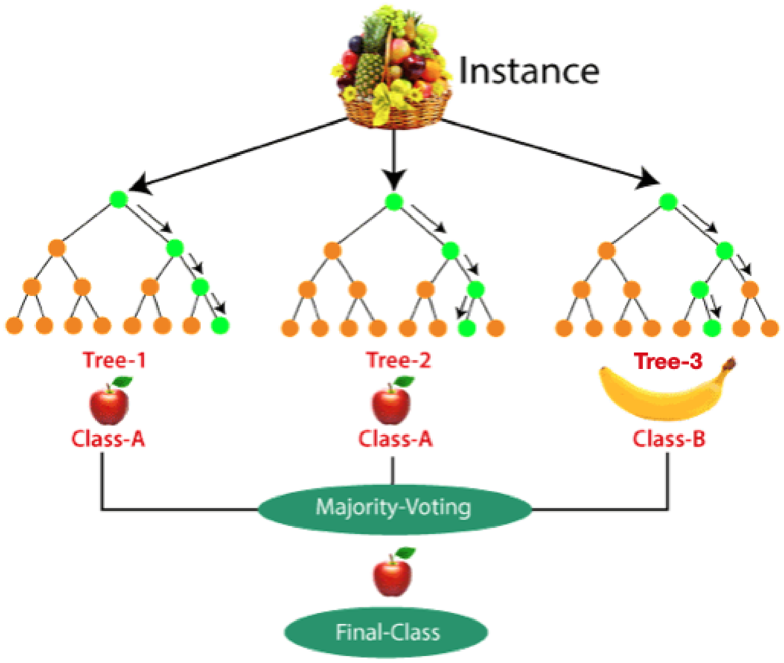
○ [0.2, 0.4)

○ [0.8, 1]

○ [0.6, 0.8)

3 1 point

Below is a depiction of a prediction being made by a random forest with $B = 3$. During the training of this random forest, there were N times a subset of features was sampled in order to determine a feature to split on. To which of the intervals below does N belong?



○ [21, 40]

○ [41, 60]

○ [1,20]

○ [81, 100]

○ [61, 80]

4 1 point

Which of the following increases model complexity (other parameters being held equal)? Select all that apply.

□ Increasing B for boosting.

□ Increasing B for random forests.

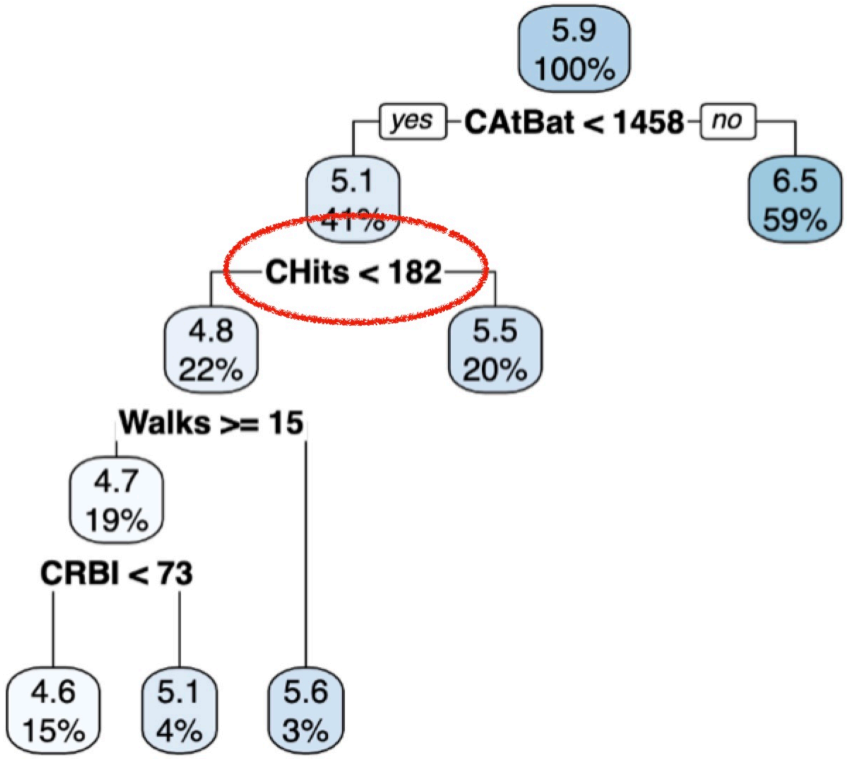
□ Increasing the number of terminal nodes for decision trees.

□ Increasing d for boosting.

□ Increasing α for decision trees.

5 1 point

During cost-complexity pruning, you remove the split circled in red. By how many nodes does this decrease the number of terminal nodes in the tree?



○ 2

○ 4

○ 5

○ 1

○ 3

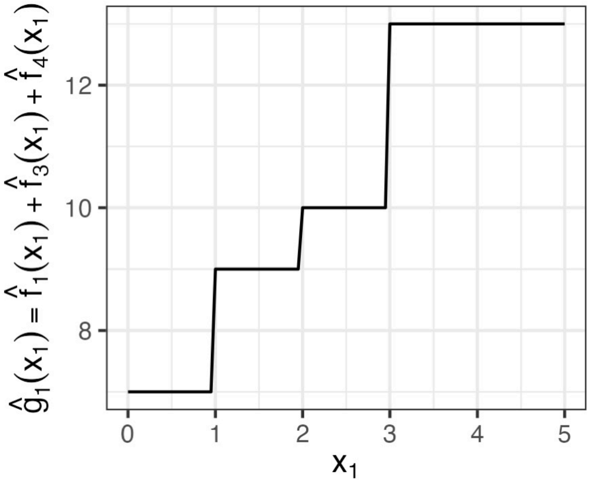
Typesetting math: 100%

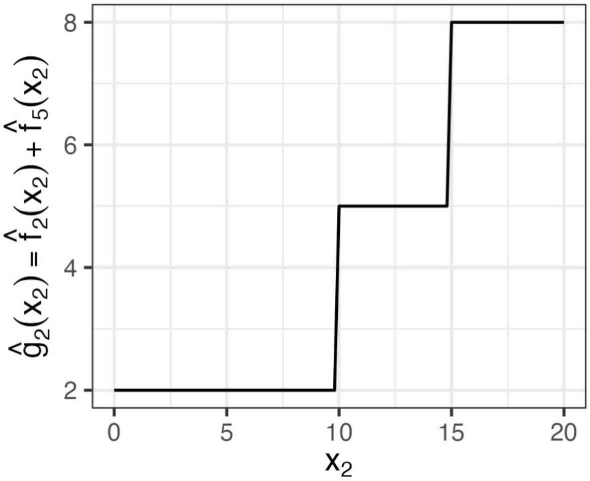
6 1 point

Suppose we have fit a boosting model with $d = 1$ and $B = 5$ to a problem with $p = 2$ features:

$$\hat{g}(x_1, x_2) = \hat{f}_1(x_1) + \hat{f}_2(x_2) + \hat{f}_3(x_1) + \hat{f}_4(x_1) + \hat{f}_5(x_2) = \hat{g}_1(x_1) + \hat{g}_2(x_2).$$

Shown below are $\hat{g}_1(x_1)$ and $\hat{g}_2(x_2)$.





When x_2 increases from 12 to 18 while x_1 stays the same, the predicted response $\hat{g}(x_1, x_2)$ increases by how much?

○ 5

○ 2

○ 3

○ 4

○ 1

7 1 point

The OOB error for random forests is most similar to which of the following quantities for the lasso?

○ CV error

○ Expected test error

○ Test error

○ Validation error

○ Training error

8

1 point

Which of the following statements are always true? Select all that apply.

□ As B increases, the variance of the random forest predictions increases.

□ As m decreases, the variance of the random forest predictions increases.

□ When m = p, a random forest reduces to a single tree fit on the training data.

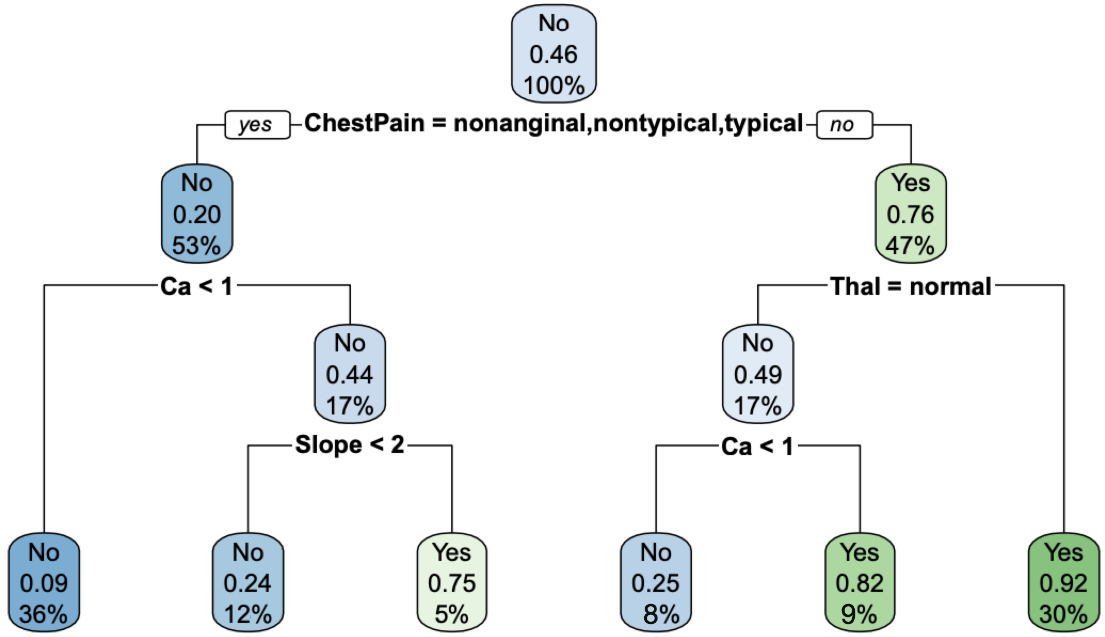
□ As m decreases, the bias of the random forest predictions increases.

□ As m decreases, the test error of the random forest decreases.

9

1 point

Consider the trained classification tree below.



Among all training observations, P percent of them have Slope < 2. To which of the intervals below does P belong?

○ [75, 100]

○ [0, 25)

○ Not enough information given.

○ [50, 75)

○ [25, 50)

Typesetting math: 100%

10 1 point

What is the maximum number of terminal nodes in a tree of interaction depth $d = 3$?

○ 8

○ 16

○ 2

○ 4

○ 6