socrative

DM-Spring-2020-Q7-Grade

18 Questions

1. (Classification) Decision Tree Decision Boundaries

2/12 A are a step-wise constant function

0/12 B I do not know

0/12 C continuous function

10/12 D are axis-parallel rectangles

2. Root Node has

10/12 A no incoming edges and zero or more outgoing edges

2/12 B one incoming edge and two or more outgoing edges

0/12 C one incoming edge and no outgoing edges

0/12 D I do not know

3. Pruning the tree means

9/12 A Simplify the tree

3/12 B Split the tree's nodes

0/12 C Merge the tree's nodes

0/12 D I do not know

4. Gini index equals to

12/12 A 1 - sum (pi^2)

0/12 B 1 + sum (pi^2)

0/12 C sum(pi * log(pi))

0/12 D -sum(pi * log(pi))

0/12 E I do not know

5. Entropy starts with 0 (rough mathematically)

3/12 A True

8/12 B False

1/12 C I do not know

6. Overall impurity measure can be obtained by

10/12 **A** a weighted average of individual rectangles

2/12 **B** majority voting

0/12 **C** I do not know

7. At each stage, we choose the split with

12/12 **A** the lowest Gini index

0/12 **B** the lowest Chi-square value

0/12 **C** the highest entropy

0/12 **D** I do not know

8. We can perform the Decision Trees in r using

11/12 **A** rpart()

0/12 **B** decisiontree()

0/12 **C** destree()

0/12 **D** reg.tree()

1/12 **E** I do not know

9. minsplit in R means

11/12 **A** the minimum number of observations that must exist in a node in order for a split to be attempted

0/12 **B** the minimum number of observations in any terminal node

1/12 **C** the minimum number of splits

0/12 **D** I do not know

10. Bagging is a technique used to reduce

10/12 **A** the variance of our predictions

1/12 **B** the bias of our predictions

1/12 **C** both

0/12 **D** I do not know

11. Bootstrap aggregation allows sampling

11/12 **A** with replacement

0/12 **B** without replacement

0/12 **C** I do not know

1/12 **D** both

12. How can Ensemble methods be constructed?

0/12 **A** By manipulating the training set

0/12 **B** By manipulating the input features

0/12 **C** By manipulating the class labels

0/12 **D** By manipulating the learning algorithm

11/12 **E** All of them

0/12 **F** None

1/12 **G** I do not know

13. Repeatedly sampling observations are taken

0/12 **A** from general population

11/12 **B** original sample data set

0/12 **C** I do not know

1/12 **D** None

14. Random Forest differs from bagging

10/12 **A** by a random sample of m predictors

2/12 **B** by bootstrapped training samples

0/12 **C** by adaptive sampling

0/12 **D** I do not know

15. Boosting differs from bagging

1/12 **A** by a random sample of m predictors

3/12 **B** by bootstrapped training samples

7/12 **C** by adaptive sampling

1/12 **D** I do not know

16. Averaging many highly correlated quantities

4/12 **A** lead to as large of a reduction in variance

5/12 **B** does not lead to as large of a reduction in variance

3/12 **C** lead to as large of a reduction in bias

0/12 **D** I do not know

17. We can perform a Random forest in R using the function

12/12 **A** randomForest()

0/12 **B** rf()

0/12 **C** randomF()

0/12 **D** boot()

0/12 **E** I do not know

18. Random Forest works

1/12 **A** for classification

0/12 **B** for regression

11/12 **C** both

0/12 **D** I do not know