



Preventing Overfitting in Decision Trees

11 试题

1
point

1.

(True/False) When learning decision trees, smaller depth USUALLY translates to lower training error.

- ☐ True
- ☒ False

1
point

2.

(True/False) If no two data points have the same input values, we can always learn a decision tree that achieves 0 training error.

- ☒ True
- ☐ False

1
point

3.

(True/False) If decision tree T1 has lower training error than decision tree T2, then T1 will always have better test error than T2.

- ☐ True

☐ False

1
point

4.
Which of the following is true for decision trees?

- ☐ Model complexity increases with size of the data.
 - ☒ Model complexity increases with depth.
 - ☐ None of the above
-

1
point

5.
Pruning and early stopping in decision trees is used to

- ☒ combat overfitting
 - ☐ improve training error
 - ☐ None of the above
-

1
point

6.
Which of the following is NOT an early stopping method?

- ☐ Stop when the tree hits a certain depth
- ☐ Stop when node has too few data points (minimum node "size")
- ☒ Stop when every possible split results in the same amount of error reduction



Stop when best split results in too small of an error reduction

1
point

7.

Consider decision tree T1 learned with minimum node size parameter = 1000. Now consider decision tree T2 trained on the same dataset and parameters, except that the minimum node size parameter is now 100. Which of the following is always true?



The depth of T2 \geq the depth of T1



The number of nodes in T2 \geq the number of nodes in T1



The test error of T2 \leq the test error of T1



The training error of T2 \leq the training error of T1

1
point

8.

Questions 8 to 11 refer to the following common scenario:

Imagine we are training a decision tree, and we are at a node. Each data point is (x_1, x_2, y) , where x_1, x_2 are features, and y is the label. The data at this node is:

x_1	x_2	y
0	1	+1
1	0	+1
0	1	+1
1	1	-1

What is the classification error at this node (assuming a majority class classifier)?

1
point

Refer to the scenario presented in Question 8.

9. If we split on x_1 , what is the classification error?

1
point

Refer to the scenario presented in Question 8.

10. If we split on x_2 , what is the classification error?

1
point

11.

Refer to the scenario presented in Question 8.

If our parameter for minimum gain in error reduction is 0.1, do we split or stop early?

- ☒ Split
- ☐ Stop early

提交测试

