Decision Trees

11 试题

1 point

1.
Questions 1 to 6 refer to the following common scenario:

Consider the following dataset:

x1	x2	x3	y
1	1	1	+1
0	1	0	-1
1	0	1	-1
0	0	1	+1

Let us train a decision tree with this data. Let's call this tree T1.
What feature will we split on at the root?

- x1
- x2
- x3

1 point

2.
Refer to the dataset presented in Question 1 to answer the following.

Fully train T1 (until each leaf has data points of the same output label). What is the depth of T1?

3

1 point

3.
Refer to the dataset presented in Question 1 to answer the following.

What is the training error of T1?

0

1 point

4.
Refer to the dataset presented in Question 1 to answer the following.

Now consider a tree T2, which splits on x1 at the root, and splits on x2 in the 1st level, and has leaves at the 2nd level. Note: this is the XOR function on features 1 and 2. What is the depth of T2?

2

1 point

5.
Refer to the dataset presented in Question 1 to answer the following.

What is the training error of T2?

0

---

1 point

6.
Refer to the dataset presented in Question 1 to answer the following.

Which has smaller depth, T1 or T2?

- ○ T1
- ● T2

---

1 point

7.
Imagine we are training a decision tree, and we are at a node. Each data point is (x1, x2, y), where x1,x2 are features, and y is the label. The data at this node is:

x1	x2	y
0	1	+1
1	0	-1
0	1	+1
1	1	+1

Which feature results in the best split?

- ○ x1
- ● x2

1 point

8.
If you are learning a decision tree, and you are at a node in which all of its data has the same y value, you should

- ( ) find the best feature to split on
- (•) create a leaf that predicts the y value of all the data
- ( ) terminate recursions on all branches and return the current tree
- ( ) go back to the PARENT node and select a DIFFERENT feature to split on so that the y values are not all the same at THIS node

---

1 point

9.
Consider two datasets D1 and D2, where D2 has the same data points as D1, but has an extra feature for each data point. Let T1 be the decision tree trained with D1, and T2 be the tree trained with D2. Which of the following is true?

- ( ) T2 has better training error than T1
- ( ) T2 has better test error than T1
- (•) Too little information to guarantee anything

---

1 point

10.
Which of these rules is more appropriate for splitting on real-valued features?

- (•) Split using thresholds (e.g., income < 60k or income >=

60k)

- ○ Split using numeric values (e.g., income == 60k, or income != 60k)
- ○ Neither of the above is appropriate

---

1 point

11.
(True/False) Decision stumps (depth 1 decision trees) are always linear classifiers.

- ● True
- ○ False

☑ I, **伟臣 沈**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.
了解荣誉准则的更多信息

提交测试