

Decision Trees in Practice

14 questions

- 1 point
1. Given an intermediate node with 6 safe loans and 3 risky loans, if the `min_node_size` parameter is 10, what should the tree learning algorithm do next?

- Create a leaf and return it
 Continue building the tree by finding the best splitting feature

- 1 point
2. Assume an intermediate node has 6 safe loans and 3 risky loans. For each of 4 possible features to split on, the error reduction is 0.0, 0.05, 0.1, and 0.14, respectively. If the minimum gain in error reduction parameter is set to 0.2, what should the tree learning algorithm do next?

- Create a leaf and return it
 Continue building the tree by using the splitting feature that gives 0.14 error reduction

- 1 point
3. Consider the prediction path `validation_set[0]` with `my_decision_tree_old` and `my_decision_tree_new`. For `my_decision_tree_new` trained with

```
1 max_depth = 6, min_node_size = 100, min_error_reduction=0.0
```

is the prediction path shorter, longer, or the same as the prediction path using `my_decision_tree_old` that ignored the early stopping conditions 2 and 3?

- Shorter
 Longer
 The same

- 1 point
4. Consider the prediction path for **ANY** new data point. For `my_decision_tree_new` trained with

```
1 max_depth = 6, min_node_size = 100, min_error_reduction=0.0
```

is the prediction path for a data point always shorter, always longer, always the same, shorter or the same, or longer or the same as for `my_decision_tree_old` that ignored the early stopping conditions 2 and 3?

- Always shorter
 Always longer
 Always the same
 Shorter or the same
 Longer or the same

- 1 point
5. For a tree trained on any dataset using parameters

```
1 max_depth = 6, min_node_size = 100, min_error_reduction=0.0
```

what is the maximum possible number of splits encountered while making a single prediction?

3



- 1 point
6. Is the validation error of the new decision tree using early stopping conditions 2 and 3) lower than, higher than, or the same as that of the old decision tree from the previous assignment?

- Higher than
 Lower than
 The same

- 1 point
7. Which tree has the smallest error on the validation data?

- model_1
 model_2
 model_3

- 1 point
8. Does the tree with the smallest error in the training data also have the smallest error in the validation data?

- Yes
 No

- 1 point
9. Is it always true that the tree with the lowest classification error on the training set will result in the lowest classification error in the validation set?

- Yes, this is ALWAYS true.
 No, this is NOT ALWAYS true.

- 1 point
10. Which tree has the largest complexity?

- model_1
 model_2
 model_3

- 1 point
11. Is it always true that the most complex tree will result in the lowest classification error in the validation set?

- Yes, this is always true.
 No, this is not always true.

- 1 point
12. Using the complexity definition, which model (model_4, model_5, or model_6) has the largest complexity?

- model_4
 model_5
 model_6

- 1 point
13. model_4 and model_5 have similar classification error on the validation set but model_5 has lower complexity. Should you pick model_5 over model_4?

- Pick model_5 over model_4
 Pick model_4 over model_5

- 1 point
14. Using the results obtained in this section, which model (model_7, model_8, or model_9) would you choose to use?

- model_7
 model_8
 model_9

I, Vamsi Vegunta, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account. Learn more about Coursera's Honor Code

Submit Quiz