Assignment-7

The due date for submitting this assignment has passed. Due on 2019-04-17, 23:59 IST.
As per our records you have not submitted this assignment.

1) In which of the following scenario a gain ratio is preferred over Information Gain? 1 point

- When a categorical variable has very large number of categories
- When a categorical variable has very small number of categories
- Number of categories is not the reason
- None of the mentioned

No, the answer is incorrect.
Score: 0
Accepted Answers:
When a categorical variable has very large number of category

2) Which of the following is/are true about Random Forest and Gradient Boosting ensemble methods? 1 point

1. Both methods can be used for classification task
2. Random Forest is use for classification whereas Gradient Boosting is use for regression task
3. Random Forest is use for regression whereas Gradient Boosting is use for Classification task
4. Both methods can be used for regression task

1 and 2
2 and 3
2 and 4
1 and 4

No, the answer is incorrect.
Score: 0
Accepted Answers:
1 and 4

3) Given an attribute table shown below, which stores the basic information of attribute a, including the row identifier of instance row_id , values of attribute values (a) and class labels of instances c. 1 point

<table>
<thead>
<tr>
<th>row_id</th>
<th>a1</th>
<th>a2</th>
<th>a3</th>
<th>a4</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1.0</td>
<td>1.1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
</tr>
</tbody>
</table>

© 2014 NPTEL - Privacy & Terms - Honor Code - FAQs -
Which of the following attribute will first provide the pure subset?

- Humidity
- Wind
- Outlook
- None of the mentioned

No, the answer is incorrect.
Score: 0
Accepted Answers: Outlook

4) True or False?
Bagging provides an averaging over a set of possible datasets, removing noisy and non-stable parts of models.

- True
- False

No, the answer is incorrect.
Score: 0
Accepted Answers: True

5) Hundreds of trees can be aggregated to form a Random forest model. Which of the following is true about any individual tree in Random Forest?

1. Individual tree is built on a subset of the features
2. Individual tree is built on all the features
3. Individual tree is built on a subset of observations
4. Individual tree is built on full set of observations

1 point
6) Boosting any algorithm takes into consideration the weak learners. Which of the following is the main reason behind using weak learners?

- Reason I: To prevent overfitting
- Reason II: To prevent underfitting

No, the answer is incorrect. Score: 0

Accepted Answers:
- Reason I
- Reason II

7) Suppose you are using a bagging based algorithm say a Random Forest in model building. Which of the following can be true?

1. Number of trees should be as large as possible.
2. You will have interpretability after using Random Forest.

- Only 1
- Only 2
- Both 1 and 2
- None of these

No, the answer is incorrect. Score: 0

Accepted Answers:
- Only 1

8) To apply bagging to regression trees which of the following is/are true in such case?

1. We build the N regression with N bootstrap sample
2. We take the average of N regression tree
3. Each tree has a high variance with low bias

- 1 and 2
- 2 and 3
- 1 and 3
- 1,2 and 3

No, the answer is incorrect. Score: 0

Accepted Answers:
- 1,2 and 3