Assignment 7

The case study for submitting this assignment has been received. All work should be submitted as indicated in the assignment guidelines. Your submission should include:

1. For the ROC curve of True positive rate vs False positive rate, which of the following are true?  
   - True
   - False
   - Can't be determined
   
   Accepted Answers: True

2. For large datasets, we should always be choosing logistic regression for classification. True or False?  
   - True
   - False
   - Can't be determined
   
   Accepted Answers: False

3. For the given confusion matrix, compute the recall.
   
<table>
<thead>
<tr>
<th>Predicted Positive</th>
<th>Predicted Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Positive</td>
<td>0.70</td>
</tr>
<tr>
<td>False Positive</td>
<td>0.60</td>
</tr>
<tr>
<td>False Negative</td>
<td>0.70</td>
</tr>
</tbody>
</table>
   
   - Recall of the dataset
   - Can't be determined
   
   Accepted Answers: Can't be determined

4. A diagnostic test is applied to 100 patients to detect a disorder, with a prevalence of 0.25% in the population. Assuming that the test has a specificity of 0.99, how many patients were incorrectly identified as having a disorder?  
   - 100
   - 9
   - Can't be determined
   
   Accepted Answers: 9

5. Do you think that the bagging procedure can be used to improve the performance of linear regression?  
   - Yes
   - No
   - Not enough information to answer the question
   
   Accepted Answers: Yes

6. The ROC plots of three classifiers are shown in the following figure.

   ![ROC plot](image)

   Which of the following statements is true about the classifiers:
   - Classifier A is better
   - Classifier B is better
   - Classifier C is better
   - Can't be determined
   
   Accepted Answers: Can't be determined

7. In which approach is the classification based on data sets whose distribution are modified in comparison to the distribution of the original training data set?
   - Bagging
   - Boosting
   - Both
   - Other
   
   Accepted Answers: All

8. Which of the following steps have been described in the diagram?

   - Bagging reduces variance of the classifier
   - Bagging increases bias of the classifier
   - Bagging helps make robust classifiers from separate classifiers
   - Bagging reduces bias of the classifier
   
   Accepted Answers: All

9. The ROC curve for a classifier A is shown in the following figure.

   ![ROC curve](image)

   - What is the True Positive Rate (TPR) at the point where the FPR is 0.1?
   
   Accepted Answers: 0.9

10. Considering the Additive algorithm, among the following statements are false?

   - In each stage, we try to fix a classifier that makes accurate predictions on a subset of the data points where the subset is a square of half the size of the data set.
   - In each stage, we try to fix a classifier that makes accurate predictions on a subset of the data points where the subset contains more of the data points where the subset contains more of the data points where the subset contains more of the data points.
   - The weight assigned to an individual classifier depends on the number of data points correctly classified by the classifier.
   - The weight assigned to an individual classifier depends on the weighted sum of the best-performing features for that classifier.
   - Can't be determined
   
   Accepted Answers: All

11. Which of the following are true:

   - Each stage is to train a classifier which makes accurate predictions on each subset of the data sets whose subset size is 1/2 the size of the data set.
   - Each stage is to train a classifier which makes accurate predictions on each subset of the data sets whose subset size is 1/2 the size of the data set.
   - Each stage is to train a classifier which makes accurate predictions on each subset of the data sets whose subset size is 1/2 the size of the data set.
   - Can't be determined
   
   Accepted Answers: All

12. Which of the following measures best evaluate the performance of a classifier?

   - Precision
   - Recall
   - Accuracy
   - Time complexity
   - Depend on the application
   
   Accepted Answers: All

DUE: 2020-03-19, 23:59 IST.