

## Unit 12 - Week 9

## Assignment 9

The due date for submitting this assignment has passed.  
As per our records you have not submitted this assignment.

**Due on 2019-10-02, 23:59 IST.**

1) Which of the following distance metric can be used in k-NN?

1 point

- Manhattan
- Jaccard
- Mahalanobis
- Minkowski
- All of the Above

No, the answer is incorrect.  
Score: 0

Accepted Answers:  
All of the Above

2) Which of the following will be true about k in k-NN in terms of Bias and Variance?

1 point

- With increase in k bias will increase and variance will not change.
- With increase in k bias will increase and variance will increase
- Bias will increase with increase in k and variance will increase with decrease in k
- None of the above

No, the answer is incorrect.  
Score: 0

Accepted Answers:  
Bias will increase with increase in k and variance will increase with decrease in k

3) Stopping Criteria for Binary Decision Tree would be

1 point

- Leaf nodes are pure
- Maximum depth is reached
- A performance metric is achieved
- All of the above

No, the answer is incorrect.  
Score: 0

Accepted Answers:  
All of the above

4) Bootstrap Sampling has been performed on a given dataset. Let say number of bootstrap sampling is  $n$  and  $n$  is being very large number. What fraction of dataset is contained in each bootstrap sample approximately?

1 point

- $\frac{1}{n}$
- $1 - \frac{1}{n}$
- $\frac{1}{3}$
- $\frac{2}{3}$

No, the answer is incorrect.  
Score: 0

Accepted Answers:  
 $\frac{2}{3}$

5) Consider the following statements:

1 point

**Statement 1 (S1):** Random forest reduces the correlation between trees  
**Statement 2 (S2):** For every tree at every split, a random subset of features are considered.

- S1 and S2 are correct and S2 is correct explanation of S1
- S1 and S2 are correct and S2 is not the correct explanation of S1
- S1 is correct and S2 is incorrect
- S1 is in correct and S2 is correct

No, the answer is incorrect.  
Score: 0

Accepted Answers:  
S1 and S2 are correct and S2 is correct explanation of S1

6) Which of the following statements is/are True?

1 point

- Model variance is high in Binary Decision Tree as compare to bagging
- Model variance is low in Binary Decision Tree as compare to bagging
- Model variance is high in Binary Decision Tree as compare to boosting
- None of the above statements are true

No, the answer is incorrect.  
Score: 0

Accepted Answers:  
Model variance is high in Binary Decision Tree as compare to bagging  
Model variance is high in Binary Decision Tree as compare to boosting

7) Why is estimating the performance of a classifier on a single dataset is not sufficient?

1 point

- Variance of parameter estimated will be high
- Misclassification error estimated will be high
- If the test data is only from one part of the distribution, we do not get the proper estimate of the performance of the classifier
- All of the above

No, the answer is incorrect.  
Score: 0

Accepted Answers:  
All of the above

8) Regression tree typically predict a constant value over a region of the input space so as to minimize the square error. If we try to minimize the absolute error. What will the filled value in the region

1 point

- Mean of the data in the region of the input space
- Mode of the data in the region of the input space
- Median of the data in the region of the input space
- None of the above

No, the answer is incorrect.  
Score: 0

Accepted Answers:  
Median of the data in the region of the input space

9) Under what conditions does fitting a linear function in the leaf nodes of a regression tree instead of the constant function not increase the variance of the model ?

1 point

- When the function is globally linear and each leaf node has at least 2 data points
- When the function is locally linear and each leaf node has at least 2 data points
- There is no change in model Variance
- None of the above

No, the answer is incorrect.  
Score: 0

Accepted Answers:  
When the function is globally linear and each leaf node has at least 2 data points

10) Which of the following is true with respect to gradient boosting?

1 point

- The weak learner typically used is a decision tree
- Any differentiable function can be used as a loss for gradient boosting
- It is necessary to use a least squares loss in the case of gradient boosting for regression problems
- Any classifier/regression model can be used as a weak learner

No, the answer is incorrect.  
Score: 0

Accepted Answers:  
The weak learner typically used is a decision tree  
Any differentiable function can be used as a loss for gradient boosting

## Course outline

How to access the portal?

Prerequisites Assignment

Matlab and Learning Modules

Week 1

Week 2

Week 3

Week 4

Week 5

Week 6

Week 7

Week 8

Week 9

- Introduction
- Knn
- Binary decision trees
- Binary regression trees
- Bagging
- Random Forest
- Boosting
- Gradient boosting
- Unsupervised learning & Kmeans
- Agglomerative clustering

Quiz : Assignment 9

Machine Learning for Engineering and Science Applications : Week 9 Feedback

Assignment 9 Solutions

Week 10

Week 11

Week 12

DOWNLOAD VIDEOS

Live Sessions

Text Transcripts