Unit 4 - Week 2

Assignment 2

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

Due on 2019-08-21, 23:59 IST.

1) Which of the following is true for a decision tree? *2 points*

a. A decision tree is an example of a linear classifier.
b. The entropy of a node typically decreases as we go down a decision tree.
c. Entropy is a measure of purity.
d. An attribute with lower mutual information should be preferred to other attributes.

a
b
c
d

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

2) Suppose there are five instances, I1, I2, I3, I4, I5 in a dataset having three features, X, Y and Z as shown in the table below: *2 points*

Instances	X	Y	Z
I1	1.6	2.3	5.1
I2	2.4	2.5	4.6
I3	3.9	3.6	3.7
I4	4.1	3.7	2.5
I5	5.6	8.3	1.8

In order to find the dependence between two variables we use the Pearson's Correlation Coefficient. Based on your understanding of Correlation Coefficient , choose the correct option/s:

a. A strong positive correlation between X and Y
b. A strong negative correlation between X and Y.
c. A weak positive correlation between X and Z.
d. A weak negative correlation between X and Z.

a
b
c
d

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

3) Suppose, you got a situation where you find that your linear regression model is under fitting the data. In such situation which of the following options would you consider? *2 points*

a. You will add more features
b. You will start introducing higher degree features
c. You will remove some features
d. None of the above.

a
b
c
d

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

4) Consider the dataset, S given below: *2 points*

Elevation	Road Type	Speed Limit	Speed
steep	Uneven	Yes	Slow
steep	Smooth	Yes	Slow
flat	Uneven	No	Fast
steep	Smooth	No	Fast

Elevation, Road Type and speed Limit are the features and Speed is the target label that we want to predict.

Find the entropy of the dataset, S as given above:

a. 0.5
b. 0
c. 1
d. 0.7

a
b
c
d

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

5) Find the information Gain if the dataset is split at the feature "Elevation": *2 points*

a. 1
b. 0
c. 0.675
d. 0.325

a
b
c
d

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

6) Find the feature on which the parent node must be chosen to split the dataset, S based on information gain: *2 points*

a. Speed Limit
b. Road Type
c. Elevation

a
b
c

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

7) Consider a simple linear regression model with One independent variable (X). The output variable is Y. The equation is : Y=aX+b where a is the slope and b is the intercept. If we change the input variable (X) by 1 unit, by how much output variable (Y) will change? *2 points*

a. 1 unit
b. By slope
c. By intercept
d. None

a
b
c
d

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

8) The following table shows the results of a recently conducted study on the correlation of the number of hours spent driving with the risk of developing acute backache. Find the equation of the best fit line for this data. *2 points*

Number of hours spent driving (x)	Risk score on a scale of 0-100 (y)
10	95
9	80
2	10
15	50
10	45
16	98
11	38
16	93

Choose which of the options is correct?

a. y = 3.39x + 11.62
b. Y = 4.69x + 12.58
c. Y = 4.59x + 12.58
d. Y = 3.59x + 10.58

a
b
c
d

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

9) You have generated data from a 3-degree polynomial with some noise. What do you expect of the model that was trained on this data using a 5-degree polynomial as function class? *2 points*

a. Low bias, high variance
b. High bias, low variance.
c. Low bias, low variance.
d. High bias, low variance.

a
b
c
d

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

10) Pruning is a technique that reduces the size of decision trees by removing sections of the tree that provide little power to classify instances. This is done in order to avoid: *2 points*

a. overfitting
b. underfitting

a
b

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