

MSITIDSCT

Edit Quiz

SAVE AND EXIT

Enable Sharing

SOC-41632384

IDSCT-2

☐

Align quiz to standard

#1

EDIT

Which of the following statements is true about outliers in Linear regression?

ANSWER CHOICE

A

Linear regression is sensitive to outliers

B

Linear regression is not sensitive to outliers

C

Can't say

D

None of these

#2

EDIT

Overfitting is more likely when you have a huge amount of data to train (Linear Regression)?

ANSWER CHOICE

A

True

B

False

C

Can't Judge

D

It depends on data

#3

EDIT

Generally, which of the following method(s) is used for predicting continuous dependent variable?

1. Linear Regression

2. Logistic Regression

ANSWER CHOICE

A

1 and 2

B

only 1

C

only 2

D

None of these

#4

EDIT

In KNN Regression, Complex concepts can be learned by local approximation using simple procedures

ANSWER CHOICE

A

True

B

False

C

Can't Judge

D

Depends on data

#5

EDIT

Which of the following are true for KNN Regression?

ANSWER CHOICE

A

There is description of the learned concepts

B


The cost of the learning process is zero


D


All the above


#6

 EDIT









You have given the following 2 statements, find which of these options is/are true in case of k-NN?

1. In case of very large value of k, we may include points from other classes into the neighborhood.

2. In case of too small value of k the algorithm is very sensitive to noise

ANSWER CHOICE

- A

1
- B


2
- C


1 and 2
- D


None of these


#7

 EDIT









Regularisation is a form of regression, in which we involve all the predictors to predict the response.

ANSWER CHOICE

- A

True
- B

False
- C

Can't predict



#8

 EDIT







In order to avoid overfitting, we can

ANSWER CHOICE

- A

Reduce number of features
- B


Keep all the features, but reduce the magnitude of the coefficients
- C


None of the above
- D


All of the above


#9

 EDIT









Which of the following statements are true?

ANSWER CHOICE

- A

Lasso regression adds penalty equivalent to absolute value of the magnitude of coefficients.
- B


Ridge regression adds penalty equivalent to the square of the magnitude of coefficients.
- C


Lasso regression adds penalty equivalent to mean value of the magnitude of coefficients.
- D


Ridge regression adds penalty equivalent to absolute value of the magnitude of coefficients.


#10

 EDIT









The logistic regression model is

ANSWER CHOICE


- A





Is a regression model
- B

a generalized linear model
- C

contains binomial errors and link logit

#11

 EDIT



Which of the following methods do we use to best fit the data in Logistic Regression?

ANSWER CHOICE

A

Least Square Error

B

Maximum Likelihood


C





Ridge loss function

D

None of these

#12

 EDIT



Which of the following is true regarding the logistic function for any value “x”?

Logistic(x): is a logistic function of any number “x”

Logit(x): is a logit function of any number “x”

Logit_inv(x): is an inverse logit function of any number “x”

ANSWER CHOICE

A

$\text{Logistic}(x) = \text{Logit}(x)$

B

$\text{Logistic}(x) = \text{Logit_inv}(x)$

C

$\text{Logit_inv}(x) = \text{Logit}(x)$

D

None of these

#13

EDIT

↑

↓

Which of the following algorithms are non-parametric models?

1. Linear regression

2. KNN regression

ANSWER CHOICE

A

1 and 2

B

1 only

C

2 only

D

None

#14

EDIT

↑

↓

Which of the following problems are regression problems?

1. Predicting the amount of view a YouTube video will get next week based on video length, the date it was posted, previous number of views, etc.

2. Predicting whether or not a Netflix user will like a particular movie.

3. Predicting the expected cab fare in New York City based on time of year, location of pickup, weather conditions etc.

ANSWER CHOICE

A

All of the three

B

Only 1

C

Only 2

D

Only 3

E


Only 1 and 3


F


Only 1 and 2


#15

 EDIT









Which of the following are important to distinct Regression and Classification problems?

ANSWER CHOICE

- A

Each type of problem may require its own specialized algorithms
- B


Each type of problem may require its own loss function to quantify how well a model performs
- C


Each type of problem may require its own specialized data collection process
- D


The data for classification problem can have categorical variables while regression problems can have only numerical data


#16

 EDIT









X	1	2	3	4	5	6
Y	10	8	15	6	7	13

X and Y values for a KNN regression are given in the following table

Calculate the predicted Y using kNN for k = 2 and MSE.

ANSWER CHOICE

- A

11.5, 12.5,7,11,9.5,6.5 and 26.67
- B


12.5, 11.5,7,11,9.5,9.5 and 22.33
- C


11.5, 12.5,7,11,9.5,6.5 and 22.33
- D


12.5, 12.5,7,11,9.5,9.5 and 26.67


#17

 EDIT









Why doesn't the predictions made by linear regression $Y = f(X) + \epsilon$ not match the observations exactly?

ANSWER CHOICE

- A

f is not a linear function
- B


the difference between prediction and observation is due to the noise term in $Y = f(X) + \epsilon$.
- C


No mathematics technique is available to compute the linear function f exactly
- D


Computational errors of linear function f


#18

 EDIT









For multi linear regression which of the following reduce standard error in parameters β_i ?

ANSWER CHOICE

- A

More data (more number of observations)
- B


Largest coverage (high variance of x)
- C


Better data (low standard deviation of the noise)
- D


More number of independent variables


#19

 EDIT









For a multilinear regression, which predictors have a substantial association with the response?

ANSWER CHOICE

- A

Predictors with small p value and large F value
- B

Predictors with large p value and large F value
- C

Predictors with large p value and small F value
- D

Predictors with small p value and small F value



#20

 EDIT







If a qualitative predictor takes 4 values in a linear regression data how many dummy variables to be created?

ANSWER CHOICE

- A

3
- B


4
- C


5
- D


2
- E

1

questions

 MULTIPLE CHOICE

 TRUE / FALSE

 SHORT ANSWER