MACHINE LEARNING

In Q1 to Q11, only one option is correct, choose the correct option:

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

A) Least Square Error B) Maximum Likelihood
C) Logarithmic Loss D) Both A and B

Correct answer-A) Least Square Error

2. Which of the following statement is true about outliers in linear regression?

A) Linear regression is sensitive to outliers B) linear regression is not sensitive to outliers
C) Can't say D) none of these

Correct answer-A) Linear regression is sensitive to outliers

3. A line falls from left to right if a slope is _____?

A) Positive B) Negative
C) Zero D) Undefined

Correct answer-B) Negative

4. Which of the following will have symmetric relation between dependent variable and independent variable?

A) Regression B) Correlation
C) Both of them D) None of these

Correct answer-B) Correlation

5. Which of the following is the reason for over fitting condition?

A) High bias and high variance B) Low bias and low variance
C) Low bias and high variance D) none of these

Correct answer-B) Low bias and low variance

6. If output involves label then that model is called as:
A) Descriptive model B) Predictive modal
C) Reinforcement learning D) All of the above

Correct answer-B) Predictive modal

7. Lasso and Ridge regression techniques belong to _____?
A) Cross validation B) Removing outliers
C) SMOTE D) Regularization

Correct answer-D) Regularization

8. To overcome with imbalance dataset which technique can be used?
A) Cross validation B) Regularization
C) Kernel D) SMOTE

Correct answer-A) Cross validation

9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?
A) TPR and FPR B) Sensitivity and precision
C) Sensitivity and Specificity D) Recall and precision

Correct answer-A) TPR and FPR

10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.
A) True B) False

Correct answer-B) False

11. Pick the feature extraction from below:
A) Construction bag of words from an email
B) Apply PCA to project high dimensional data
C) Removing stop words
D) Forward selection

Correct answer-B) Apply PCA to project high dimensional data

In Q12, more than one options are correct, choose all the correct options:

12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?
A) We don't have to choose the learning rate
B) It becomes slow when number of features is very large.
C) We need to iterate.
D) It does not make use of dependent variable.

Correct answer-
A) We don't have to choose the learning rate
B) It becomes slow when number of features is very large.
D) It does not make use of dependent variable.

Q13 and Q15 are subjective answer type questions, Answer them briefly.

13. Explain the term regularization?

Answer-
When we use regression models to train some data, there is a good chance that the model will over fit the given training data set. Regularization helps sort this over fitting problem by restricting the degree of freedom of a given equation that is simply reducing the number of degrees of a polynomial function by reducing their corresponding weights.

In a linear equation, we do not have huge weight/coefficients as a small change in weight can make a large difference for the dependent variable(Y). So regularization constraints the weights of such features to avoid over fitting. to regularize the model, a Shrinkage penalty is added to the function.

14. Which particular algorithms are used for regularization?

Answer-
We use Lasso regularization to overcome the problem that ridge has, Lasso(Least Absolute Shrinkage and Selection Operator) is an alternative that can pick relevant features that will be useful for modeling.

Lasso also has the shrinkage parameter but the difference that has with Ridge is that there is no squared term of the estimated coefficient but only an absolute value.

To overcome the problem that ridge has, Lasso(Least Absolute Shrinkage and Selection Operator) is an alternative that can pick relevant features that will be useful for modeling. Lasso also has the shrinkage parameter but the difference that has with Ridge is that there is no squared term of the estimated coefficient but only an absolute value. the term after RSS is called the shrinkage penalty or L1 norm.

15. Explain the term error present in linear regression equation?

Answer-
An error term is a residual variable produced by a statistical or mathematical model, which is created when the model does not fully represent the actual relationship between the independent variables and the dependent variables. As a result of this incomplete relationship, the error term is the amount at which the equation may differ during empirical analysis.

The error term is also known as the residual, disturbance, or remainder term, and is variously represented in models by the letters e, ε, or u.

--------//----------//----------