MACHINE LEARNING (*SOLUTIONS*)

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

Solution: (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

Solution: (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

Solution: (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

Solution: (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

Solution: (C) Low bias and high 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

Solution: (B) Predictive modal

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

Solution: (D) Regularization

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

Solution: (D) SMOTE

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

Solution: (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

Solution: (B) False

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

Solution: (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.

Solution:
(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.

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

13. Explain the term regularization?

Answer:
Regularization: Regularization is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting.

14. Which particular algorithms are used for regularization?

Answer:
The commonly used regularization techniques (Algorithms) are:
(1) L1 regularization (Lasso regression)
(2) L2 regularization (Ridge regression)
(3) Dropout regularization (Dilution regularization)

L1 regularization:
A regression model which uses L1 Regularization technique is called LASSO (Least Absolute Shrinkage and Selection Operator) regression. L1 regularization gives output in binary weights from 0 to 1 for the model's features and is adopted for decreasing the number of features in a huge dimensional dataset.
Lasso Regression adds *"absolute value of magnitude"* of coefficient as penalty term to the loss function (L).

L2 regularization:
A regression model that uses L2 regularization technique is called Ridge regression.
L2 regularization disperse the error terms in all the weights that leads to more accurate customized final models.
Ridge regression adds "*squared magnitude*" of coefficient as penalty term to the loss function (L).

Dropout regularization:
Dropout (also called Dilution) is a regularization technique for reducing overfitting in artificial neural networks by preventing complex co-adaptations on training data. It is an efficient way of performing model averaging with neural networks.

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.

- An error term appears in a statistical model, like a regression model, to indicate the uncertainty in the model.
- The error term is a residual variable that accounts for a lack of perfect goodness of fit.
- Heteroskedastic refers to a condition in which the variance of the residual term, or error term, in a regression model varies widely.

Error Term Use in a Formula:
An error term essentially means that the model is not completely accurate and results in differing results during real-world applications. For example, assume there is a multiple linear regression function that takes the following form:

$$Y=\alpha X+\beta\rho+\epsilon$$

where:
α,β=Constant parameters
X,ρ=Independent variables
ε=Error term
When the actual Y differs from the expected or predicted Y in the model during an empirical test, then the error term does not equal 0, which means there are other factors that influence Y.