

## MACHINE LEARNING

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

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

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

Answer : A) positive

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

Answer : C) Both of them

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

Answer : 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

Answer : B) Predictive modal

7. Lasso and Ridge regression techniques belong to \_\_\_\_\_?

A) Cross validation B) Removing outliers C) SMOTE D) Regularization

Answer : D) Regularization

8. To overcome with imbalance dataset which technique can be used?

A) Cross validation B) Regularization C) Kernel D) SMOTE

Answer : B) Regularization

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

Answer : A) TPR and FR

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

A) True B) False

Answer : A) True

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

Answer : B) Apply PCA to project high dimensional data

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.

ASSIGNMENT – 39 MACHINE LEARNING Q13 and Q15 are subjective answer type questions, Answer them briefly.

13. Explain the term regularization?

Answer : To make things regular or acceptable. This is exactly why we use it for. Regularization are techniques used to reduce the error by fitting a function appropriately on the given training data and avoid overfitting.

14. Which particular algorithms are used for regularization?

Answer : LASSO, Ridge and Elastic-Net regression.

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

Answer : A Regression line always has an error term because, in real life, independent variables are never perfect predictors of the dependent variable. Rather the line is an estimate based on the variable data. So the error term tells you how certain you can be about the formula.