

# 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
- 

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
- 

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

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

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
- 

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
-

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
- 

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

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

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

- A) Cross validation                      B) Regularization  
C) Kernel                                  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
- 

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

- A) True                      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
- 

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.
- 

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

### 13. Explain the term regularization?

**Ans:** Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting. In the context of machine learning, regularization is the process which regularizes or shrinks the coefficients towards zero. In simple words, regularization discourages learning a more complex or flexible model, to prevent overfitting. Based on Occam's Razor, Regularization is one of the key concepts in Machine learning. It helps prevent the problem of overfitting, makes the model more robust, and decreases the complexity of a model. Regularization works by adding a penalty or complexity term or shrinkage term with Residual Sum of Squares (RSS) to the complex model.  $\beta_0, \beta_1, \dots$  ... In simple linear regression, our optimization function or loss function is known as the residual sum of squares (RSS).

There are two types of regularization as follows:

L1 Regularization or Lasso Regularization.

L2 Regularization or Ridge Regularization.

L1 regularization penalizes the sum of absolute values of the weights, whereas L2 regularization penalizes the sum of squares of the weights.

---

### 14. Which particular algorithms are used for regularization?

**Ans:** There are three main regularization techniques, namely:

Ridge Regression (L2 Norm)

Lasso (L1 Norm)

Dropout

Ridge and Lasso can be used for any algorithms involving weight parameters, including neural nets. Dropout is primarily used in any kind of neural networks e.g. ANN, DNN, CNN or RNN to moderate the learning. Let's take a closer look at each of the techniques.

---

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

**Ans:** For a very simple example, suppose you are predicting the weight of adult human males based on their height. Well, height is certainly related to weight - taller people tend to be heavier - but the model won't be perfect because there is a range of weights at each height. The error is the difference between the predicted value and the actual value.

For a very simple example, suppose you are predicting the weight of adult human males based on their height. Well, height is certainly related to weight - taller people tend to be heavier - but the model won't be perfect because there is a range of weights at each height. The error is the difference between the predicted value and the actual value.

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

**N.B:-In MCQ question right answer marked by yellow colour.**