# 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

2. Which of the following statement is true about outliers in linear regression?
A) Linear regression is sensitive to outliers

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

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

5. Which of the following is the reason for over fitting condition?
A) High bias and high variance

6. If output involves label then that model is called as:
B) Predictive modal

7. Lasso and Ridge regression techniques belong to _____?
D) Regularization

8. To overcome with imbalance dataset which technique can be used?
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
C) Sensitivity and Specificity

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

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?
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?
Ans: - Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting.

14. Which particular algorithms are used for regularization?

Ans:- LASSO L1 FORM

- RIDGE L2 FORM

- ELASTICNET(LESS POPULAR) combination of both L-1 and L-2

  Lasso (L-1 form) and Ridge (L-2 form) act like a features selection

if match the relationship between the each feature and label if in any feature and label relationship it don't find +ve effect it don't consider them while prediction.

- L-1 don't give any importance to those features who don't have any relation with label
- L-2 give very little importance to those features who don't have any relation with label doesn't act like proper feature selection

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

  Ans:- Error term is the term in a model regression equation that tallies up and accounts for the unexplained difference between the actually observed values of the independent variable and the results predicted by the model. Hence, the error term is a measure of how accurately the regression model reflects the actual relationship between the independent and dependent variable or variables.