Machine learning assignment

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 _____?
A) Positive

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?
C) Low bias and high variance

6. If output involves label then that model is called as:
A) Descriptive model

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:
B) Apply PCA to project high dimensional data

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.

13. When a regression model is overfitted, regularization helps to sort the problem. We use regularization to penalize the model to learn at a slower pace so that it will learn better.

14. For regularization different algorithms are used. They are

i.LASSO(L1 form)

ii.RIDGE(L2 form)

15. An error term means that the model is not completely accurate and differs in results during real world applications.

Let, the multiple linear regression equation is given by,

$Y=aX+bZ+err$

Where, a,b are constant parameters

X,Z are independent variables

Err=error term

When the actual Y differs from the expected or predicted Y in the model, then the error term does not equal to 0, which means there are other factors that influence Y.