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 _____?
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?
C) Low bias and high variance

6. If output involves label then that model is called as:
D) All of the above

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

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

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?
c)We need to iterate

13)REGULARIZATION:
Regularization is the technique used to reduce the data, and to avoid the overfitting, to make our model in high accuiracy….

14)LASSO and RIDGE are the most commonly used algorithm used in Regularization.
LASSO(Least Absolute Shrinkage and Selection Operator)also call as L1 form,It is also acts as an features selection which will completely ignore the unwanted features.
RIDGE is also call as an L2 form but here it will not completely ignore,it will consider but gives very less importance…

15) The error term is known as the residual, disturbance, or remainder term .An error is a variable in 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.