MACHINE LEARNING ASSIGNMENT

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

Ans. Least square error

2. Which of the following statement is true about outliers in linear regression?

Ans. Linear regression is sensitive to outliers

3. A line falls from left to right if a slope is Negative?
4. Which of the following will have symmetric relation between dependent variable and independent variable?

Ans. Coorelation

5. Which of the following is the reason for over fitting condition?

Ans. Low bias and high variance

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

Ans. Predictive model

7. Lasso and Ridge regression techniques belong to _____?

Ans. Regularization

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

Ans. Cross validation

9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

Ans. FPR and TPR

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

Ans. False

11. . Pick the feature extraction from below

Ans. Apply PCA to high dimensional data

12. 2. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

Ans. a. we don't have to choose the learning rate
b. it becomes slow when number of features is very large.

13. Explain the term regularization?

ans. When we use regression model to train some data, there is some good chance that the

model will overfit the given training data set. Regularization helps sort this overfitting problem by restricting the degrees of freedom of a given equation .i.e., simply reducing the number of degrees of polynomial function by reducing their corresponding weights.

14. . Which particular algorithms are used for regularization?

Ans. LASSO and ridge algorithms are used for regularization.

LASSO (Least absolute shrinkage and selection operator) : LASSO regression penalizes the model based on the sum of the magnitude of the co-efficient.

Ridge regression : Ridge regression penalizes the model based on the sum of the square of the magnitude of the co-efficient.

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

Ans. The error term in linear regression equation is given by actual value minus predicted value.

There are error term which find the errors:

1. Mean absolute error (MAE) : it represents average error
2. Mean squared error (MSE) : similar to MAE but noise is exaggerated and larger errors are punished. It is harder to interpret than MAE as its not in base units, however it is generally more popular.
3. Root mean squared error (RMSE) : most popular metric, similar to MAE, however the result is squared rooted to make it more interpretableas its in base units.