MACHINE LEARNING

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

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

13. Explain the term regularization?

- It is a technique used to reduce the errors by fitting the functions appropriately on a given set.
- It is used to calibrate ML models in order to minimize the loss and to prevent the overfitting and under fitting.
- There are 2 types of regularisation L1 & L2
- It is used to reduce the complexity of the model

14. Which particular algorithms are used for regularization?

- Ridge Regression
    - It shrinks the coefficients as it helps to reduce the model complexity and multi collinearity
- Lasso(least absolute shrinkage and selection operator)
    - It coverts the coefficients of less important features to zero which helps in feature selection and shrinks the coefficients of remaining features to reduce the model complexity and avoiding the overfitting.
- Elastic-Net regression
    - It is a regularised regression which combines L1 & l2 penalties of LASSO and Ridge methods respectively.

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

- An error term essentially means the model is not completely accurate and results in differing results during real world applications.
- For eg with linear regression model tracking a stock price over a time , the error term is the difference between the expected price at a particular time and the price that was actually observed.