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

Ans- A) Least Square Error

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

Ans- A) Linear regression is sensitive to outliers

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

A) Positive B) Negative

C) Zero D) Undefined

Ans- B) Negative

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

Ans- A) Regression

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

Ans- C) Low bias and high variance

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

A) Descriptive mode B) Predictive modal

C) Reinforcement learning D) Reinforcement learning

Ans- B) Predictive modal

7. Lasso and Ridge regression techniques belong to _____?

A) Cross validation B) Removing outliers

C) SMOTE D) Regularization

Ans D) Regularization

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

A) Cross validation B) Regularization

C) Kernel D) SMOTE

Ans- 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

Ans- A) TPR and FPR

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

A) True B) False

Ans- 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

Ans- 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?

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.

Ans- B) It becomes slow when number of features is very large.

13. Explain the term regularization?

Ans- Regularization is a technique used in machine learning to prevent overfitting and improve the generalization of a model. Overfitting occurs when a model fits the training data too closely, capturing noise or random fluctuations that do not represent the true underlying patterns in the data. Regularization introduces a penalty term to the model's objective function, discouraging overly complex models that may fit the training data too well.

There are different types of regularization, and two common ones are L1 regularization (Lasso) and L2 regularization (Ridge). These methods add a regularization term to the cost function, which is a function of the model parameters. The regularization term is scaled by a hyperparameter (lambda or alpha), which controls the strength of the regularization
L1 regularization adds the absolute values of the coefficients as a penalty term, encouraging sparsity (some coefficients become exactly zero). This is useful for feature selection

L2 regularization adds the squared values of the coefficients as a penalty term. It tends to distribute the weight more evenly among all features, preventing extreme values and reducing the impact of any single feature.

Regularization helps in achieving a balance between fitting the training data well and avoiding overfitting, leading to a model that performs better on unseen data.

14. Which particular algorithms are used for regularization?

Ans Regularization techniques are commonly applied to linear regression models. Two popular regularization algorithms are:

1. **Lasso Regression (L1 Regularization):** Lasso stands for Least Absolute Shrinkage and Selection Operator. In Lasso regression, the regularization term added to the cost function is the absolute sum of the coefficients multiplied by a hyperparameter (lambda or alpha). This regularization term encourages sparsity in the model, meaning that some coefficients can become exactly zero. Lasso is useful for feature selection, as it tends to eliminate less important features.
2. **Ridge Regression (L2 Regularization):** Ridge regression adds the squared sum of the coefficients to the cost function, multiplied by a hyperparameter (lambda or alpha). The regularization term in Ridge regression prevents the model from becoming too complex by penalizing large coefficients. Ridge regression is effective when there is multicollinearity (high correlation) among the predictor variables.
3. These regularization techniques help prevent overfitting by discouraging overly complex models and improving their generalization performance on new, unseen data. The choice between L1 and L2 regularization depends on the specific characteristics of the data and the goals of the modeling task.

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

Ans- In the context of linear regression, the term "error" typically refers to the difference between the predicted values of the dependent variable (or response variable) and the actual observed values. This difference is also known as the residual. The linear regression equation is generally represented as:

$$Y=\beta_0+\beta_1X_1+\beta_2X_2+...+\beta_nX_n+\epsilon$$

Y is the dependent variable (the variable we are trying to predict).

1,2,...,$X1,X2,...,Xn$ are the independent variables (features or predictors).

$0,1,\ldots,\beta_0,\beta_1,\ldots,\beta_n$ are the coefficients, representing the weights assigned to each independent variable.

ϵ is the error term, representing the difference between the predicted and actual values that the model is unable to explain.

The goal of linear regression is to find the values of the coefficients ($0, 1,\ldots, \beta_0,\beta_1,\ldots,\beta_n$) that minimize the sum of the squared errors (residuals). This process is often carried out using a method called the least squares method.

The error term (ϵ) accounts for unobserved factors or randomness in the relationship between the independent and dependent variables. In an ideal scenario, the errors are assumed to follow certain statistical properties, such as being normally distributed with a mean of zero.

By minimizing the sum of squared errors, linear regression aims to create a line (or hyperplane in multiple dimensions) that best fits the observed data points, allowing for the prediction of the dependent variable based on the values of the independent variables.