Machine Learning

Q1) Which of the following methods do we use to find the best fit line for data in Linear Regression?

Answer- A) Least Square Error.

Q2) Which of the following statement is true about outliers in linear regression?

Answer- A) Linear Regression is sensitive to outliers.

Q3) A line falls from left to right if a slope is _____?

Answer- B) Negative.

Q4) Which of the following will have symmetric relation between dependent variable and independent variable?

Answer- B) Correlation.

Q5) Which of the following is the reason for over fitting condition?

Answer- C) Low bias and high variance.

Q6) If output involves label then that model is called as:

Answer- B) Predictive model

Q7) Lasso and Ridge regression techniques belong to_____?

Answer- D) Regularization

Q8) To overcome with imbalance dataset which technique can be used?

Answer- A) Cross Validation

Q9) The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

Answer- A) TPR and FPR

Q10) In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

Answer- A)True.

Machine Learning

Q11) Pick the feature extraction from below:

Answer- A) Construction bag of words from a email

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

Answer- A) We don't have to choose the learning rate.

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

Q13) Explain the term regularization?

Answer-

Regularization is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting i.e. it helps to simply reducing the no of degrees of a polynomial function by reducing their corresponding weights.

To regularize the model, a Shrinkage penalty is added to the cost function.

Types of Regularization -

1) LASSO
2) RIDGE
3) ELASTICNET

LASSO (Least Absolute shrinkage and Selection Operator) Regression-

A regression model which uses L1 Regularization technique is called LASSO.

LASSO Regression penalizes the model based on the sum of magnitude of the coefficients.

The regularization term is given by regularization= λ *$\sum|\beta j|$

Where, λ is the shrinkage factor.

RIDGE

RIDGE Regression penalizes the model based on the sum of squares of magnitude of the coefficients.

The regularization term is given by regularization= λ *$\sum|\beta \hat{} 2j|$

Where, λ is the shrinkage factor.

Q14) Which particular algorithms are used for regularization?

Answer-

There are two commonly used algorithms in regularization-

1)L1 Regularization

L1 Regularization uses LASSO which is based on the sum of magnitude of the coefficients.

2)L2 Regularization

L2 Regularization uses RIDGE which is based on the sum of squares of magnitude of the coefficients.

Q15) Explain the term error present in linear regression equation?

Answer-

An error term represents the margin of error within a statistical model; it refers to the sum of the deviations within the regression line, which provides an explanation for the difference between the theoretical value of the model and the actual observed results. The regression line is used as a point of analysis when attempting to determine the correlation between one independent variable and one dependent variable.

$$Y=\alpha X+\beta\rho+\epsilon$$

where:

α, β =Constant parameters

X, ρ =Independent variables

ϵ =Error term