MACHINE LEARNING

In Q1 to Q11, only one option is correct, choose the correct option:

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?
   Least Square Error
2. Which of the following statement is true about outliers in linear regression?
   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?
   Regression
5. Which of the following is the reason for over fitting condition?
   Low bias and high variance
6. If output involves label then that model is called as:
   Predictive model
7. Lasso and Ridge regression techniques belong to _____?
   Regularization
8. To overcome with imbalance dataset which technique can be used?
   SMOTE
9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?
   TPR and FPR
10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.
    False
11. Pick the feature extraction from below:
    Construction bag of words from a email

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?
    We don't have to choose the learning rate, It becomes slow when number of features is very large., We need to iterate

MACHINE LEARNING

Q13 and Q15 are subjective answer type questions, Answer them briefly.

13. Explain the term regularization?

Regularization is one of the most important concepts of machine learning. is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting.

There are mainly 3 types of regularization techniques

1. L1 regularization
2. L2 regularization
3. Dropout regularization

Regression model which uses L1 Regularization technique is called LASSO(Least Absolute Shrinkage and Selection Operator) regression.

Regression model that uses L2 regularization technique is called Ridge regression.

Lasso Regression adds “absolute value of magnitude” of coefficient as penalty term to the loss function(L).

Ridge regression adds “squared magnitude” of coefficient as penalty term to the loss function(L)

14. Which particular algorithms are used for regularization?

1. Ridge Regression
2. LASSO (Least Absolute Shrinkage and Selection Operator) Regression
3. Elastic-Net Regression

Ridge regression is a method for analyzing data that suffer from multi-collinearity.

LASSO is a regression analysis method that performs both feature selection and regularization in order to enhance the prediction accuracy of the model.

Elastic-Net is a regularized regression method that linearly combines the L1 and L2 penalties of the LASSO and Ridge methods respectively.

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

Tearm Error is a residual variable produced by 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. As a result of this incomplete relationship, the error term is the amount at which the equation may differ during empirical analysis.

The error term is also known as the residual, disturbance, or remainder term, and is variously represented in models by the letters e, ε, or u.