**12-05-2023**

**Answers: - Machine Learning Assignment**

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

A1) Least Square Error (A)

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

A2) Linear regression is sensitive to outliers (A)

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

A3) Negative (B)

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

A4) Correlation (B)

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

A5) Low bias and high variance (C)

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

A6) Descriptive model (A)

Q7) Lasso and Ridge regression techniques belong to \_\_\_\_\_\_\_\_\_?

A7) Regularization (D)

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

A8) SMOTE (D)

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

A9) TPR and FPR

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

A10) False (B)

Q11) Pick the feature extraction from below:

A11) Apply PCA to project high dimensional data (B)

Q12) Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

A12) 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.0

Q13) Explain the term Regularization

A13) Regularization is a machine learning technique which prevent the model from overfitting by omitting the attributes which are not contributable towards the output or reduces the coefficient difference internally. Some times the ML model works well in training data but does not perform well in testing data. It means the model is not able to predict output when deals with unseen data, hence the model is called overfitted. This problem can be dealt with regularization.

Q14) Which particular algorithms are used for regularization?

A14) Regularization can be done through 2 ways.

1) Lasso Regression 2) Ridge Regression.

1. Lasso Regression – Model will automatically omit the no. of attributes internally which model thinks is not contributable towards Y (output)
2. Ridge Regression – Model will reduce down the coefficient differences internally.

The combination of the Lasso regression & Ridge regression is ElasticNet. Which omit the attributes and also reduces the coefficient values.

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

A15) The equation of the linear regression is Y = a +bx + ḕ. Where ḕ stands for error. The error in the equation is the mean square error. The difference between the predicted value and the actual value can be called as ḕ (error).