MACHINE LEARNING-WORKSHEET1

Ques.1-Ques.11 (choose one correct answer)

Ques1: Which of the following methods do we use to find the best fit line for data in Linear Regression?

Ans1: A) Least Square Error

Que2: Which of the following statement is true about outliers in linear regression?

Ans2: A) Linear regression is sensitive to outliers

Ques3: A line falls from left to right if a slope is _____?

Ans3: B) Negative

Ques4: Which of the following will have symmetric relation between dependent variable and independent variable?

Ans4: B) Correlation

Ques5: Which of the following is the reason of over fitting condition?

Ans5: A) High bias and high variance

Ques6: If output involves label then that model is called as:

Ans6: B) Predictive model

Ques7: Lasso and Ridge regression techniques belong to_____?

Ans7: D) Regularization

Ques8: To overcome with imbalance dataset which technique can be used?

Ans8: D) SMOTE

Ques9: The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

Ans9: A) TPR and FPR

Ques10: In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

Ans10: B) False

Ques11: Pick the feature extraction from below:

Ans11: A) Apply PCA to project high dimensional data

Ques.12 (multiple correct answer)

Ques12: Which of the following is true about Normal Equation used complete the coefficient of the Linear regression?

Ans12: A) We don't have to choose the learning rate

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

Ques.13-Ques.15

Ques13: Explain the term regularization?

Ans13: The problem of over fitting is faced in Machine Learning model so the regularization technique is used to overcome from it. The problem of over fitting is occur due to

- If there are more & more number of features whether they are necessary to use or not
- If our data is biased

So, we use regularization technique and they are of three types but now use only two of them that are Lasso regularization and Ridge regularization.

Ques14: Which particular algorithms are used for regularization?

Ans14: There are two Machine learning algorithms which are used for a regularization technique:

- LASSO: (Least Absolute Shrinkage and Selection Operator)
  LASSO is denoted by L1 and in this technique, it does not consider those features which have no contribution in predicting the data.

- RIDGE:
  Ridge Regularization is denotes by L2 and in this technique, it consider all those features in analyses whether they contribute or not in predicting the data but give a very less importance to those features.

Ques15: Explain the term error present in linear regression equation?

Ans15: The error term is also known as residual. Residual is the distance between the actual value and predicted value. The error term occurs means that the independent variables are not as good as they want in prediction of the dependent variable. As lesser the error better will be the model. The linear regression equation is:

$$Y = a + bX + e$$

Where, a = intercept of Y

b = slope of a line

e = error term or residual

To calculate residual in machine learning we have three methods which are:

- Mean Absolute Error (average of error)
- Mean Square Error (average of error and ignore the large error)
- Root Mean Square Error