Answer 1: A. Least Square Error

Answer2: A. Linear regression is sensitive to outliers

Answer3: A. Positive

Answer4: B. Correlation

Answer5: D. none of these

Answer6: D. All of the above

Answer7: D. Regularization

Answer8: D. SMOTE

Answer9: A. TPR and FPR

Answer10: A. True

Answer11: B. Apply PCA to project high dimensional data

Answer12: A&B

Answer13: Regularization –

Regularization is one of the most important concepts of machine learning. It is a technique to prevent the model from over fitting by adding extra information to it. Sometimes the machine learning model performs well with the training data but does not perform well with the test data. It means the model is not able to predict the output when deals with unseen data by introducing noise in the output, and hence the model is called over fitted. This problem can be deal with the help of a regularization technique. There are two types of regularization techniques

1. Ridge regression

2. Lasso regression

This technique can be used in such a way that it will allow to maintain all variables or features in the model by reducing the magnitude of the variables. Hence, it maintains accuracy as well as a generalization of the model. It mainly regularizes or reduces the coefficient of features toward zero. In simple words, "In regularization technique, we reduce the magnitude of the features by keeping the same number of features.

Answer 14: The algorithms used for regularization are as follows:

a. Ridge Regression - Ridge regression is a method for analysing data that suffer from multi-collinearity. Ridge regression is also known as the L2 Regularization. **Ridge regression shrinks the coefficients as it helps to reduce the model complexity and multi-collinearity.**

**b. Lasso Regression – also called least absolute shrinkage and selection operator, hence** LASSO is a regression analysis method that performs both feature selection and regularization in order to enhance the prediction accuracy of the model. LASSO regression is also known as the L1 Regularization. **LASSO regression converts coefficients of less important features to zero, which indeed helps in feature selection, and it shrinks the coefficients of remaining features to reduce the model complexity, hence avoiding over fitting**

**Answer15:**

**Error term present in linear regression equation -** An error term 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. An error term appears in a statistical model, like a regression model, to indicate the uncertainty in the model. An error term represents the margin of error within a statistical model it refers to the sum of 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.