MachineLearning_assignment

1. A) Least Square Error

2. A) Linear regression is sensitive to outliers

3. B) Negative

4. A) Regression

5. C) Low bias and high variance

6. B) Predictive modal

7. D) Regularization

8. D) SMOTE

9. A) TPR and FPR

10. B) False

11. B) Apply PCA to project high dimensional data

12. A) We don't have to choose the learning rate. , B) It becomes slow when number of features is very large.

13.ANS: Regularization is one of the most important concepts of machine learning. It is a technique to prevent the model from overfitting 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 overfitted. This problem can be deal with the help of a regularization technique. 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.

14. ANS: There are mainly two types of regularization techniques, which are given below:
**Ridge Regression**

**Lasso Regression**

<span style="color:purple">Ridge Regression:–</span>
1. Ridge regression is one of the types of linear regression in which a small amount of bias is introduced so that we can get better long-term predictions.
2. Ridge regression is a regularization technique, which is used to reduce the complexity of the model. It is also called as L2 regularization.
3. In this technique, the cost function is altered by adding the penalty term to it. The amount of bias added to the model is called Ridge Regression penalty. We can calculate it by multiplying with the lambda to the squared weight of each individual feature.

<span style="color:purple">Lasso Regression:-</span>
1. Lasso regression is another regularization technique to reduce the complexity of the model. It stands for Least **Absolute and Selection Operator.**
2. It is similar to the Ridge Regression except that the penalty term contains only the absolute weights instead of a square of weights.
3. Since it takes absolute values, hence, it can shrink the slope to 0, whereas Ridge Regression can only shrink it near to 0.
4. It is also called as L1 **regularization.** The equation for the cost function of Lasso regression will be:

15. ANS: 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 $e, \varepsilon,$ or u 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.