MACHINE LEARNING

C SOMA SEKHAR GOUD

BATCH: DS2307

1. A) Least Square Error

2. A) Linear regression is sensitive to outliers

3. B) Negative

4. B) Correlation

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

C) We need to iterate.

D) It does not make use of the dependent variable.

13.

Regularization is a technique used in machine learning to prevent overfitting and improve the generalization of a model. which refers to the modifications that can be made to a learning algorithm that helps to reduce this generalization error and not the training error. It reduces by ignoring the less important features. It also helps prevent overfitting, making the model more robust and decreasing the complexity of a model.

14.

L1 Norm (Lasso), L2 Norm (Ridge Regression) are the two most commonly used types of regularization techniques.

L1 adds the “absolute value of magnitude” of the coefficient as a penalty term to the loss function(L).

Whereas L2 adds the “squared magnitude” of the coefficient as a penalty term to the loss function(L).

15.

The term "error" refers to the difference between the actual observed values of the dependent variable (the target) and the predicted values generated by the linear regression model.

The linear regression equation (single variable) is : $y=mx+b+\varepsilon$

ε : error

These errors can happen for a bunch of reasons, like randomness playing tricks, measurements not being totally accurate, or even hidden factors that the model doesn't grasp. So, the main game in linear regression is to make these errors as small as possible- like squishing them together. We do this by finding the line that reduces the total squared errors, which means it's the best line that comes closest to our real data points. By diving into these errors, we can fine-tune our model, see how well it's doing, and get a clearer picture of how different variables are connected.