Machine Learning

1: Least Square Error

2: Linear regression is sensitive to outliers

3: Negative

4: Correlation

5: low bias and high variance

6: Predictive model

7: Regularization

8: SMOTE

9: True Positive Rate (TPR) and False Positive Rate (FPR)

10: False

11: Constructing bag of words from an email

<u>Q13 and Q15 are subjective answer type questions, Answer them briefly</u>

13: Regularization is a technique used in machine learning and statistical modeling to prevent overfitting and improve the generalization of a model by adding a penalty term to the loss function. The primary goal of regularization is to balance fitting the training data well while avoiding overly complex models that might not generalize to new, unseen data.

In the context of linear regression or other models, regularization typically involves adding a penalty term to the model's loss function.

Two common types of regularization are L1 regularization (Lasso) and L2 regularization (Ridge):

L1 Regularization (Lasso):

1. Adds a penalty term proportional to the absolute value of the coefficients (L1 norm) to the loss function.
2. Encourages sparsity by pushing some coefficients to exactly zero, effectively performing feature selection.
3. Useful when dealing with a large number of features, as it tends to eliminate less important features.

L2 Regularization (Ridge):

1. Adds a penalty term proportional to the square of the coefficients (L2 norm) to the loss function.
2. Controls the magnitude of the coefficients, keeping them small but rarely pushing them to zero.
3. Helps prevent large variations in the coefficients, thus reducing model complexity.

The regularization parameter (lambda or alpha) controls the strength of the regularization effect. Larger values of lambda or alpha result in stronger regularization, leading to more penalty on the coefficients and potentially simpler models.

Regularization helps in achieving models that generalize better to unseen data by reducing overfitting, which occurs when a model learns the noise or specific characteristics of the training data too well but fails to perform well on new data. It's particularly valuable when dealing with high-dimensional data or when the number of features exceeds the number of samples, as it helps in controlling model complexity and improving its stability.