SOLUTIONS

Answer 1. (D) Both A and B (Least Square Error and Maximum Likelihood)

Answer 2. (A) Linear regression is sensitive to outliers

Answer 3. (B) Negative

Answer 4. (D) None of these

Answer 5. (C) Low bias and High variance

Answer 6. (A) Predictive Model

Answer 7. (D) Regularization

Answer 8. (D) SMOTE

Answer 9. (A) TPR and FPR

Answer 10. (B) False

Answer 11. (A) Construction bag of words from a email

Answer 12. (A) And (D) ((A) We don't have to choose the learning rate. And (D) it does not make use of dependent variable.)

Answer 13. Regularization is a machine learning approach used to avoid overfitting, which happens when a model is overly complicated and captures noise and random oscillations in the training data, resulting in subpar generalisation performance on fresh data.

By including a penalty term in the cost function of the model, regularization prevents the model from learning intricate correlations from the training data that might not transfer well to fresh data. This penalty term often depends on the size of the model's coefficients.

There are two common types of Regularization techniques:

1. L1 (Lasso regularization techniques)
2. L2 (Ridge regularization techniques)

A penalty term proportional to the absolute value of the coefficients is added to the cost function by L1 regularisation. By successfully completing feature selection and lowering the model's complexity, this penalty term encourages the model to set some of the coefficients to zero.

On the other hand, L2 regularisation increases the cost function's penalty component, which is inversely proportional to the square of the coefficients. This penalty factor induces the model to learn smaller coefficient values, effectively reducing the coefficients to zero.

Answer 14.

L1 regularisation (Lasso Regression): This algorithm modifies the cost function by adding a penalty term proportionate to the absolute value of the model weights. It is useful for feature selection since it favours sparse solutions.

L2 regularisation (Ridge Regression): This approach adds a penalty term to the cost function that is proportional to the model weights squared. It aids in the prevention of overfitting and can increase model stability.

Early stopping: It is a technique for terminating the training process before the model begins to overfit the training data.

Data augmentation: It is a strategy that generates additional training data by applying different transformations to existing data. It can help to prevent overfitting and improve model performance.

Elastic Net: L1 and L2 regularisation are combined in the elastic net. It may be able to get beyond some of the restrictions imposed by each unique algorithm.

Dropout: This is a neural network training method that randomly removes some units. Both overfitting and generalisation can be avoided.

Answer 15.

The difference between the anticipated and actual values of the dependent variable is represented by the error term in linear regression

The linear regression equation is written as follows:

$Y = a + bX + e$

Where Y is the dependent variable, X is the independent variable, a and b are the regression equation coefficients, and e is the error term.

The error term is the unexplained variance in the dependent variable that the independent variable cannot account for. It is a measurement of the difference between the observed and expected values of the dependent variable based on the independent variable.

The error term should ideally be random and normally distributed, with a mean of zero. If the error term is not random, it may suggest that the linear regression model is inappropriate for the data or that there are other factors influencing the dependent variable that the model does not for account.