Answers for Machine Learning Worksheet

Ans1- a) Least Square Error

Ans2- a) Linear regression is sensitive to outliers

Ans3- b) Negative

Ans4- b) Correlation

Ans5- c) Low bias and high variance

Ans6- b) Predictive Model

Ans7- d) Regularization

Ans8- d) SMOTE

Ans9- a) TPR and FPR

Ans11- a) Construction bag of words from an email

Ans12- a) and b)

Ans13- Regularization is a technique used to reduce errors by fitting the function appropriately on the given training set and avoiding over fitting. The commonly used regularization techniques are:

i) Lasso Regularization
ii) Ridge Regularization
iii) Elastic Net Regularization

Ans14- There are several algorithms used for regularization in machine learning. The most common ones are Lasso (L1), Ridge (L2) and Elastic Net.

L1 regularization is used when we have a large number of features and we want to reduce the number of features. It does this by adding an L1 penalty equal to the absolute value of the magnitude of coefficients.

L2 regularization is used when we have large number of features and we want to avoid over fitting.

Elastic Net is a combination of both L1 and L2 regularization.

Ans15- The error in a regression equation is the last term in the equation. It accounts for the unexplained difference between the actually observed values of the independent variable and the results predicted by the model. For example, within the linear regression model tracking a stock's price over time, the error term is the difference between the expected price at a particular time and the price that was actually observed.