Machine learning Worksheet

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 model
7. D) Regularization
8. D) SMOTE
9. A) TPR and FPR
10. B) False
11. A) Constructing bag of words from an email
12. A) We do not have to choose the learning rate,
    B) It becomes slow when the number of features is very large

13. Regularization is a technique used in machine learning to prevent overfitting and improve the generalization of a model. It involves adding a penalty term to the loss function during model training. The penalty term discourages the model from fitting the training data too closely and encourages it to generalize better to new data. It mainly regularizes or reduces the coefficient of features toward zero. In simple words, "In regularization technique, we reduce the magnitude of the features by keeping the same number of features."

14. There are three main regularization techniques, namely:

    I. Ridge Regression (L2 Norm)
    II. Lasso (L1 Norm)
    III. Dropout

    Ridge and Lasso can be used for any algorithms involving weight parameters, including neural nets. Dropout is primarily used in any kind of neural networks e.g. ANN, DNN, CNN or RNN to moderate the learning. Let's take a closer look at each of the techniques.

- L1(Lasso) - LASSO is a regression analysis method that performs both feature selection and regularization in order to enhance the prediction accuracy of the model. It adds a penalty ***(L1 penalty)*** to the loss function that is equivalent to the magnitude of the coefficients which has the effect of forcing some of the coefficient estimates to be exactly equal to zero when the regularization parameter is sufficiently large.
- L2(Ridge) - Ridge regression is a method for analyzing data that suffer from multi-collinearity. Ridge regression adds a penalty ***(L2 penalty)*** to the loss function that is equivalent to the square of the magnitude of the coefficients. The regularization parameter *(λ)* regularizes the coefficients such that if the coefficients take large values, the loss function is penalized.

- Dropout- It is a regularization technique used in neural networks. It prevents complex co-adaptations from other neurons. In neural nets, fully connected layers are more prone to

overfitting on training data. Using dropout, you can drop connections with a 1-p probability for each of the specified layers. Where p is called the keep probability parameter and which needs to be tuned.

15. 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 explains 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. The error in the linear regression is the difference between the observed value and the predicted value of the dependent value(Y).