Machine learning Worksheet

1. D) Both Least Mean Square and Maximum Likelihood
2. A) Linear Regression is sensitive to outliers.
3. B) Negative
4. C) Both Regression and Correlation
5. C) Low Bias and High Variance
6. B) Predictive model
7. D) Regularization
8.
9. A) True Positive Rate(TPR) and False Positive Rate(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.

14. The working of the algorithms are same, the only difference is in the loss function. Algorithms used for regularization are-
- 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.

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