MACHINE LEARNING ASSIGNMENT

1. Maximum Likelihood

2. Linear regression is sensitive to outliers

3.Negative

4. Regression

5.C) Low bias and high variance

6. Predictive modal

7. Regularization

8.SMOTE

9. TPR and FPR

10. False

11.Construction bag of words from a email

12 .A) We don't have to choose the learning rate.

B) It becomes slow when number of features is very large.

13. Explain the term regularization?

ANS: Regularization is a method used to reduce or refine the errors by fitting the most suitable functions on the given training data set and avoid overfitting. Regularization help us to reduces the variance of the model, without a substantial increment in its bias.

There are three regularization techniques are used :

L1 regularization

L2 regularization

Dropout regularization

L1 REGULARIZATION: A regression model which uses L1 Regularization method is called LASSO(Least Absolute Shrinkage and Selection Operator) regression. lasso regression takes the magnitude of the coefficients

L2 REGULARIZATION :regression model that uses L2 regularization technique is called Ridge regression. It takes the square of the coefficient.

DROPOUT REGULARIZATION: Dropout is used as a regularization technique which prevents overfitting by ensuring that there is no units are codependent.

14. Which particular algorithms are used for regularization?

ANS: There are 3 type of algorithm that are used in regularization

1.Ridge Regression: Ridge regression is a method for analyzing data that affected from multi-collinearity.It takes the square of the coefficient. Ridge regression is also known as the L2 Regularization.

2.LASSO (Least Absolute Shrinkage and Selection Operator) Regression:LASSO is a regression analysis method that is make to performs both feature selection and regularization in order to increase the prediction accuracy of a model.LASSO regression is better known as the L1 Regularization . lasso regression takes the magnitude of the coefficients

3.Elastic-Net Regression:Elastic-Net is a regularized regression method that is linearly combines the L1 and L2 penalties of the LASSO and Ridge methods respectively.

15.Explain the term error present in linear regression equation?

An error term is the term in a model regression equation which tallies up and accounts for the uncertain difference between the actually observed values of the independent variable and the results predicted by the model. Thus the error term is a measure of how accurate the regression model reflects the actual relationship between the independent and dependent variable or variables. The error term can indicate either that the model can be improved, such as by adding in another independent variable that explains some or all of the difference, or

adding randomness, meaning that the dependent and independent variable or variables are not correlated to any greater degree.Error and residual are almost meaning.