MACHINE LEARNING

1. Which of the following methods do we use to find the best fit line for data in Linear Regression? **A) Least Square Error**
2. Which of the following statement is true about outliers in linear regression? **A) Linear regression is sensitive to outliers**
3. A line falls from left to right if a slope is **B) Negative**
4. **(C) Both Regression and correlation** will have symmetric relation between dependent variable and independent variable.
5. Which of the following is the reason for over fitting condition **c) Low bias and high variance**
6. If output involves label then that model is called **(A) Descriptive model**
7. Lasso and Ridge regression techniques belong to (**D) Regularization**
8. To overcome the imbalance dataset **(D)** SMOTE technique is used
9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses **(C) sensitivity and specificity** to make graph
10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less **False**
11. A) Construction bag of words from a email B) Apply PCA to project high dimensional data C) Removing stop words D) Forward selection – **All the above are the features of extraction**
12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression? **A) We don't have to choose the learning rate. B) It becomes slow when number of features is very large**
13. **REGULARISATION: R**egularisation is a technique in machine learning which is used to reduce over fitting. It leads to smoothening of

the regression line and thus prevents over fitting by penalising the loss function.

14. **ALGORITHMS USED FOR REGULARISATION :**
    - Lasso regression
    - Ridge regression
15. **ERROR:** the term error is the difference between the actual value and the predicted value. The vertical distance between the data point and the regression line is known as error or residual. Each data point has one residual and the sum of all the differences is known as the Sum of Residuals/Errors.