MACHINE LEARNING WORKSHEET

1. **Which of the following methods do we use to find the best fit line for data in Linear Regression?**
   A) Least Square Error B) Maximum Likelihood C) Logarithmic Loss D) Both A and B
   Ans. Least squared method

2. **Which of the following statement is true about outliers in linear regression? A) Linear regression is sensitive to outliers B) linear regression is not sensitive to outliers C) Can't say D) none of these**
   Ans, A) Sensitive to outliers

3. **A line falls from left to right if a slope is _____? A) Positive B) Negative C) Zero D) Undefined**
   Ans.B) Negative

4. **Which of the following will have symmetric relation between dependent variable and independent variable? A) Regression B) Correlation C) Both of them D) None of these**
   Ans. B) Correlation

5. **Which of the following is the reason for over fitting condition? A) High bias and high variance B) Low bias and low variance C) Low bias and high variance D) none of these.**
   Ans.C) Low bias and high variance
   Since an overfitting model is trained too much, there is less bias however when the model deals with an unseen dataset, it is unable to generalize thus producing a lot of error in prediction thus giving us a high variance.

6. **If output involves label then that model is called as: A) Descriptive model B) Predictive model C) Reinforcement learning D) All of the above**
   Ans. B) Predictive model

7. **Lasso and Ridge regression techniques belong to _____? A) Cross validation B) Removing outliers C) SMOTE D) Regularization**
   Ans. C) Regularization
   These are regularization techniques which improves the quality of prediction of a model by avoiding overfitting/ underfitting of a model.

8. **To overcome with imbalance dataset which technique can be used? A) Cross validation B) Regularization C) Kernel D) SMOTE**
   Ans. D)SMOTE
   SMOTE works by oversampling minority instances to overcome imbalance dataset

9. **The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph? A) TPR and FPR B) Sensitivity and precision C) Sensitivity and Specificity D) Recall and precision**
   Ans. A) TPR and FPR

10. **In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less. A) True B) False**
    Ans. False

The AUC ROC curve must have more area under the curve for a better model as it gets better at distinguishing between the positive and negative classes.

11. **Pick the feature extraction from below: A) Construction bag of words from a email B) Apply PCA to project high dimensional data C) Removing stop words D) Forward selection**

    Ans A) Construction bag of words from a email

    In Q12, more than one options are correct, choose all the correct options:

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. C) We need to iterate. D) It does not make use of dependent variable.**

    Ans A) We don't have to choose the learning rate. B) It becomes slow when number of features is very large. C) We need to iterate.

    As Normal Equation does make use of dependent variable, option D is not selected while the rest are true.

13. **Explain the term regularization?**

    Ans. Regularization:

    A good machine learning model is expected to generalize the data well between the training data and unseen data and give us a reliable output. The performance of a model can reduce due overfitting or underfitting. In an overfitting model, the model is run through too many iterations to achieve minimum global error, however while doing so, the model starts to also fit noisy data from the dataset which is unnecessary for predicting the optimum output thus ignoring the trend seen in the dataset. Hence, it is unable to generalize when fed with a new dataset giving us a poor predicted value. Conversely, an underfitting model also performs poorly as the algorithm has not trained through enough data and fails to capture the correct trend.

    One of the technique to avoid overfitting or underfitting of a model is through regularization. Regularization is a method in which it either eliminates or shrinks the weightage of those coefficients that would contribute to a high cost function by adding penalty to its loss function. Few examples of regularization are L1(LASSO) and L2(RIDGE) regularization techniques which are used in linear regression models.

14. **Which particular algorithms are used for regularization?**

    Ans. The most common regularization algorithms are discussed below:

    LASSO (Least Absolute Shrinkage and Selection Operator: In the case of LASSO or L1 regularization, the feature that contributes least will be eliminated or given zero importance while predicting the label.

    Ridge regression or L2 regularizes by shrinking the weightage of the coefficient closer to zero and giving that particular feature to contribute the least while predicting the label.

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

    Ans. The difference between the expected value and predicted value is the error term present in a linear equation. The equation is given below:

$$Y = \alpha X + \beta \rho + \epsilon$$

where:α,β=Constant parameters X,ρ=Independent variables ϵ=Error term