Machine Learning worksheet 1

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 1) – A- Least Square error

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 2) - 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 3)- 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 4)- B- Corelation

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 5)- C- Low Bias and high variance

6. If output involves label then that model is called as:

A) Descriptive model **B) Predictive modal**

C) Reinforcement learning **D) All of the above**

Ans 6)-B- Predictive Modal

7. Lasso and Ridge regression techniques belong to _____?

A) Cross validation **B) Removing outliers**

C) SMOTE **D) Regularization**

Ans7)-D- Regularization

8. To overcome with imbalance dataset which technique can be used?

A) Cross validation **B) Regularization**

C) Kernel **D) SMOTE**

Ans 8)-D- SMOTE

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 9)- 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--

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 11)- B- Apply PCA to project high dimensional data

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------

13. **Explain the term regularization?**

Ans- Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting.

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

Ans- Regularization algorithms like LASSO, Ridge, and Elastic-Net regression.

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

Ans- 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 provides an explanation for 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.