**My Name –Vaishali Ramdas Salve**

**Date- 26/12/2023**

**Batch No – DS2311**

**Sir – Shwetank Mishra**

**Assignment Name- Machine Learning**

**In Q1 to Q11, only one option is correct option:**

1. Which of the following method do we use to find the best fit line for data in linear Regression?

D) Both A and B

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. which of the following will have symmetric relation between dependent variable and independent variable?

B) Correlation

5. which of the following is the reason for over fitting condition?

c) Low bias and high variance

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

B) Predictive model

7. Lasso and Ridge regression technique belong to ------?

D) Regularization

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

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

10. In AUC Receiver Operator Characteristics ( AUCROC ) curve for the better model area under the curve should be less.

B) False

11. Pick the feature extraction from below:

A) Constructing bag of words from an email

**In Q12, more than one option 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 learning rate.

**Machine Learning**

Q13 and Q15 are subjective answer types question , Answer them briefly.

13. Explain the term Regularization?

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 cost function , discouraging a cost on large coefficients. This helps to control the magnitude of the coefficient and, in turn, reduces the risk of the mpdel fitting the training data too closely, making it more robust and applicable to unseen data.

14. which particular algorithms are used for regularization ?

Several algorithms incorporate regularization techniques to prevent overfitting and enhance model generalization techniques to prevent overfitting and enhance model generalization (losso) and L2 regularization (Ridge). These techniques are commonly applies to linear regression and logistic regression model. Additionally , the Elastic Net algorithm combines both L1 and L2 regularization.

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

The term “error” in the context of a linear regression equation refers to the difference between the actual abserved values and the values predictes by the linear regression model.in equation:

***Y=B0+B1X +E***

. Y is the dependent variable,

. X is the independent variable,

. B0 is the intercept,

. B1 is the slope,

. E represents the error term.

The error term E captures the unodserved factors or random variations that influence the dependent variable but are not accountant for by the liner relationship with the independent variable. The goal of linear regression is to minimize these error, often using the method of last squares, to find the best – fitting line that describes the relationship between the variables.

b