MACHINE LEARNING – WORKSHEET 1: ANSWERS

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
   Answer: 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
   Answer: A) Linear regression is sensitive to outliers

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

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

Answer: C) Low bias and 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
   Answer: B) Predictive model

7. Lasso and Ridge regression techniques belong to _____?
   A) Cross validation
   B) Removing outliers
   C) SMOTE
   D) Regularization
   E) **Answer:** D) Regularization

8. To overcome with imbalance dataset which technique can be used?
   A) Cross validation
   B) Regularization
   C) Kernel
   D) SMOTE
   E) **Answer:** 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
   Answer: 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
    Answer: True

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
    Answer: 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
    Answer: A) We don't have to choose the learning rate and B) It becomes slow when number of features is very large

13. Explain the term regularization?
    Regularization refers to techniques that are used to calibrate machine learning models in order to minimize the adjusted loss function and prevent overfitting or underfitting. Using Regularization, we can fit our machine learning model appropriately on a given test set and hence reduce the errors in it.

14. Which particular algorithms are used for regularization?
    The most common types of regularization algorithms are lasso, ridge, and elastic net.
    - **Lasso Regularization**: Lasso is a type of regularization that uses L1-norm Regularization. Lasso regularization works by adding a penalty to the absolute value of the magnitude of coefficients. This forces certain coefficients to be equal to zero, which in turn helps to reduce overfitting.

    - **Ridge Regularization**: Ridge is a type of regularization that uses L2-norm Regularization. Ridge regularization works by adding a penalty to the square of the magnitude of coefficients. This forces all coefficients to be close to zero but does not allow them to be equal to zero. It **is** effective at reducing overfitting.

    - **Elastic net Regularization**: Elastic net is a type of regularization that combines L1-norm Regularization and L2-norm Regularization. It is effective at both reducing overfitting and improving interpretability.

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

An error term is a residual variable produced by a statistical or mathematical model, which is created when the model does not fully represent the actual relationship between the independent variables and the dependent variables. The error term is also known as the residual. Residual is the distance between the actual Y and the predicted Y, as shown below:
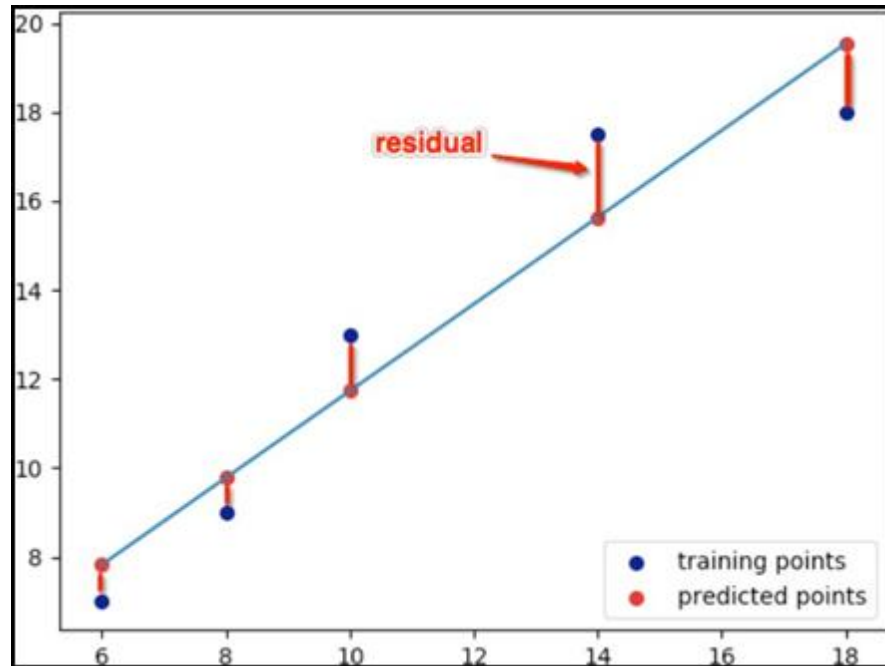
Mathematically, Residual is:

$$r = y - (mx + b)$$

Hence, the sum of the square of residuals is:

$$r_i = y_i - (mx_i + b) \qquad \text{(Residual for one point)}$$

$$\sum_{i=1}^{n} r_i = \sum_{i=1}^{n} (y_i - (mx_i + b)) \qquad \text{(Sum of residuals)}$$

$$R(x) = \sum_{i}^{n} r_i^2 = \sum_{}^{n} (y_i - (mx_i + b))^2 \qquad \text{(Sum of squares of residuals)}$$

** ** **