MACHINE LEARNING ASSIGNMENT

In Q1 to Q11, only one option is correct, choose the correct option:

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: 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
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
Answer: D) Regularization

8. To overcome with imbalance dataset which technique can be used?
   A) Cross validation
   B) Regularization
   C) Kernel

D) SMOTE
Answer: A) Cross validation

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: B) False

11. Pick the feature extraction from below:
    A) Construction bag of words from an 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

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.
Answer: A), B) and C)

Q13 and Q15 are subjective answer type questions, Answer them briefly.

13. Explain the term regularization?

Regularization in machine learning refers to a set of techniques used by data scientists to prevent overfitting. It refers to the modifications that can be made to a learning algorithm that helps to reduce this generalization error (a measure of how accurately an algorithm can predict outcome values for previously unseen data) and not the training error. It reduces by ignoring the less important features. It also helps prevent overfitting, making the model more robust and decreasing the complexity of a model.

It also means restricting a model to avoid overfitting by shrinking the coefficient estimates to zero. When a model suffers from overfitting, we should control the model's complexity. Technically, regularization avoids overfitting by adding a penalty to the model's loss function:

Regularization = Loss function + penalty

14. Which particular algorithms are used for Regularization?

There are three common techniques used for Regularization:
- L1 Regularization
- L2 Regularization
- Elastic Net

L1 Regularization: It is also called as LASSO (Least Absolute Shrinkage and Selection Operator) regression. In L1 regularization it ignores the least important features and emphasize more on model's essential features, which makes some coefficients zero.
it modifies the over-fitted or under-fitted models by adding the penalty equivalent to the sum of the absolute values of coefficients. it takes the true values of coefficients.
This means that the coefficient sum can also be 0, because of the presence of negative coefficients. The cost function for Lasso regression:

$$\text{Cost function} = \text{Loss} + \lambda \times \sum\|w\|$$

L2 Regularization: It is also called as Ridge regression. it modifies the over-fitted or under fitted models by adding the penalty equivalent to the sum of the squares of the magnitude of coefficients. The magnitude of coefficients is squared and added. Ridge Regression performs regularization by shrinking the coefficients present.
The L2 regularization technique tries to keep the model's weights close to zero, but not zero, which means each feature should have a low impact on the output while the model's accuracy should be as high as possible. The cost function of ridge regression:

$$\text{Cost function} = \text{Loss} + \lambda \times \sum\|w\|^2$$

Elastic net : it is the combination of L1 and L2 regularization.

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

An error present in linear Regression equation is the difference between the theoretical value and the actual observed value of the model. It represents the margin of error within the statistical model. The regression line is used as a point of analysis while determining the correlation between the independent and dependent variable. it also refers to the sum of deviations within the regression line. The error of the regression provides the absolute measure of the typical distance that the data points fall from the regression line.