MACHINE LEARNING

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

   ANS: 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: 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

   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

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: B) Predictive model

7. Lasso and Ridge regression techniques belong to_____?
   A) Cross validation B) Removing outliers
   C) SMOTE D) Regularization

   ANS: D) Regularization

MACHINE LEARNING

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

   ANS: 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: 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: A) 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

    ANS: 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.

    ANS: A), B) and C)

MACHINE LEARNING

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

13. Explain the term regularization?

Ans: Regularization is used for preventing overfitting and improve the generalization ability of a model in ML. When a ML model is trained on a dataset, it may become too complex and start memorizing the training data instead of learning the underlying patterns. So, the model may perform badly on new, unseen data.

Regularization helps address this issue by adding a penalty term to the model's objective function during training. This penalty term discourages the model from fitting the training data too closely and instead encourages it to find simpler, more generalizable patterns.

14. Which particular algorithms are used for regularization?

Ans: Particular algorithms are used for regularization in ML:

a. **Lasso (L1 Regularization):** In Lasso, the penalty term added to the objective function is proportional to the absolute values of the model's coefficients. This technique encourages sparse feature selection, driving some coefficients to zero and effectively performing feature selection. Lasso can be useful when dealing with high-dimensional datasets where only a few features are relevant.

b. **Ridge (L2 Regularization):** Ridge adds a penalty term proportional to the squared magnitudes of the model's coefficients. This technique encourages the model to distribute the importance of the features more evenly and reduces the impact of individual features. Ridge is commonly used and helps to stabilize the model by reducing the variance.

c. **Elastic Net Regularization:** Elastic Net combines Lasso and Ridge by adding both penalty terms to the objective function. This technique addresses some of the limitations of Lasso and Ridge and provides a balance between feature selection (Lasso) and coefficient shrinkage (Ridge).

d. **Dropout**: Dropout is a regularization technique used in neural networks. During training, dropout randomly sets a fraction of the output activations (neurons) to zero at each update, effectively "dropping out" those neurons. This prevents the model from relying too heavily on a specific subset of neurons and encourages it to learn more robust features.

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

Ans: "Error" in Linear regression refers to the dissimilarity between the predicted values of the dependent variable in linear regression model and the actual observed values of the dependent variable in the dataset. It is also known as the "residual" or "prediction error."

$y = \beta_0 + \beta_1 x + \varepsilon$

where:

y is the dependent variable (the variable we are trying to predict),

x is the independent variable (the variable used to predict y),

β_0 is the intercept (the y-value when x is zero),

β_1 is the slope (the change in y for a unit change in x), and

ε represents the error term.