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

   Answer::(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 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::(D) None of these

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 modal
   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

MACHINE LEARNING

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

   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::(B)False

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

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, D)

MACHINE LEARNING

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

13. Explain the term regularization?

Regularization is a technique used in machine learning to prevent overfitting and improve the generalization performance of a model.

Overfitting occurs when a model learns the noise or random fluctuations in the training data, resulting in poor performance on new, unseen data.

The main goal of regularization is to strike a balance between fitting the training data well and avoiding excessive complexity in the model. .

Regularization helps address this issue by adding a regularization term to the loss function, which penalizes complex or large parameter values.

The most commonly used regularization techniques in machine learning are L1 regularization (Lasso) and L2 regularization (Ridge). L1 regularization adds a penalty term proportional to the absolute value of the coefficients, encouraging sparsity and feature selection. L2 regularization adds a penalty term proportional to the square of the coefficients, which encourages small and well-distributed coefficient values.

Regularization techniques are widely used in various machine learning algorithms, including linear regression, logistic regression, support vector machines, and neural networks. The choice of regularization technique and the strength of the regularization term (controlled by a hyperparameter) depends on the specific problem, dataset, and the desired trade-off between model complexity and performance.

14. Which particular algorithms are used for regularization?

Several algorithms incorporate regularization techniques to mitigate overfitting and improve generalization performance. Some commonly used algorithms that utilize regularization include:

1. **Ridge Regression:** Ridge regression is a linear regression model that incorporates L2 regularization. It adds a penalty term proportional to the squared magnitude of the coefficients to the loss function, encouraging smaller and more evenly distributed coefficients.

2. **Lasso Regression**: Lasso regression is another form of linear regression that employs L1 regularization. It adds a penalty term proportional to the absolute value of the coefficients, which promotes sparsity and feature selection by driving some coefficients to zero.

3. **Elastic Net:** Elastic Net is a regression model that combines L1 and L2 regularization. It adds a penalty term that is a combination of the L1 and L2 norms of the coefficients. Elastic Net can be beneficial when there are many correlated features in the dataset.

4. **Support Vector Machines (SVM):** SVMs can use regularization through the C parameter, which controls the trade-off between achieving a low training error and minimizing the norm of the model's weight vector. A higher C value results in less regularization, potentially leading to a higher training accuracy but potentially overfitting the data.

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

In the context of linear regression, the term "error" refers to the discrepancy between the actual observed values and the predicted values produced by the linear regression model. It represents the unexplained or residual variation in the dependent variable that cannot be accounted for by the linear relationship with the independent variables.

MACHINE LEARNING

The linear regression equation is typically represented as:

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p + \varepsilon$

In this equation:

- y represents the dependent variable or the target variable.

- $x_1, x_2, ..., x_p$ represent the independent variables or predictors.

- $\beta_0, \beta_1, \beta_2, ..., \beta_p$ are the coefficients or parameters estimated by the regression model.

- ε represents the error term.

The error term, ε, captures the difference between the actual observed values of the dependent variable and the predicted values obtained from the linear regression equation

. It accounts for all the factors that affect the dependent variable but are not captured by the linear relationship with the independent variables. These factors could include measurement errors, unobserved variables, random noise, or any other sources of variability that cannot be explained by the model.

The goal of linear regression is to estimate the coefficients ($\beta_0, \beta_1, \beta_2, ..., \beta_p$) that minimize the sum of squared errors (SSE) or the sum of squared residuals (SSR). By minimizing the errors, the regression model aims to find the best-fit line or hyperplane that represents the linear relationship between the independent variables and the dependent variable.

Overall**, the error term in linear regression represents the discrepancy between the observed values and the predicted values, and minimizing these errors is a key objective in estimating the parameters of the linear regression model.**