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

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- 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

MACHINE LEARNING

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.

- **A) We don't have to choose the learning rate.**
- **B) It becomes slow when number of features is very large.**

MACHINE LEARNING

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

13. Explain the term regularization?

Ans- Regularization is a technique used in statistical models and machine learning to prevent overfitting, which occurs when a model learns noise in the training data rather than the underlying patterns. Overfitting can lead to poor performance on unseen data.

Here's how regularization works:

1. **Penalty Term**: Regularization adds a penalty to the loss function that the model minimizes. This penalty discourages overly complex models by constraining the size of the coefficients.
2. **Types of Regularization**:
    - **Lasso (L1 Regularization)**: Adds the absolute values of the coefficients as a penalty. This can lead to some coefficients being exactly zero, effectively performing feature selection.
    - **Ridge (L2 Regularization)**: Adds the squared values of the coefficients as a penalty. It tends to shrink the coefficients but does not set any to zero, maintaining all features in the model.
    - **Elastic Net**: Combines both L1 and L2 penalties, balancing between the two.
3. **Impact on Model**: By including regularization, the model becomes simpler and more generalizable, leading to better performance on unseen data. It helps to control the trade-off between fitting the training data well and maintaining model simplicity.
4. **Tuning**: The strength of the regularization effect is controlled by a hyperparameter (often denoted as λ or alpha), which can be tuned using cross-validation to find the optimal value.

14. Which particular algorithms are used for regularization?

Ans- Several algorithms incorporate regularization techniques to improve model performance and prevent overfitting. Here are some of the most common ones:

1. **Linear Regression**:
    - **Ridge Regression (L2 Regularization)**: Adds a penalty equal to the square of the magnitude of coefficients.
    - **Lasso Regression (L1 Regularization)**: Adds a penalty equal to the absolute value of the coefficients, potentially driving some coefficients to zero.
    - **Elastic Net**: Combines both L1 and L2 penalties.
2. **Logistic Regression**:
    - Similar to linear regression, logistic regression can use L1 (Lasso), L2 (Ridge), or Elastic Net regularization to improve classification performance.
3. **Support Vector Machines (SVM)**:
    - SVMs can include regularization parameters that control the trade-off between achieving a low error on the training data and maintaining a simple model.
4. **Neural Networks**:
    - **Weight Decay**: Often implemented as L2 regularization.
    - **Dropout**: A technique where a fraction of neurons are randomly ignored during training, helping to prevent overfitting.
5. **Decision Trees and Ensemble Methods**:
    - **Random Forests**: Introduce randomness in feature selection and can include constraints

MACHINE LEARNING

like maximum tree depth.

- **Gradient Boosting Machines (GBM)**: Can include regularization techniques in their loss functions, such as L1 or L2 penalties on leaf weights.

6. **Generalized Additive Models (GAMs)**:
   - Can incorporate penalties on the smoothness of the functions used to estimate relationships.
7. **K-Nearest Neighbors (KNN)**:
   - While KNN itself doesn't directly use regularization, techniques like feature selection and scaling can be considered forms of regularization in the broader sense.

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

Ans- Several algorithms incorporate regularization techniques to improve model performance and prevent overfitting. Here are some of the most common ones:

1. **Linear Regression**:
   - **Ridge Regression (L2 Regularization)**: Adds a penalty equal to the square of the magnitude of coefficients.
   - **Lasso Regression (L1 Regularization)**: Adds a penalty equal to the absolute value of the coefficients, potentially driving some coefficients to zero.
   - **Elastic Net**: Combines both L1 and L2 penalties.
2. **Logistic Regression**:
   - Similar to linear regression, logistic regression can use L1 (Lasso), L2 (Ridge), or Elastic Net regularization to improve classification performance.
3. **Support Vector Machines (SVM)**:
   - SVMs can include regularization parameters that control the trade-off between achieving a low error on the training data and maintaining a simple model.
4. **Neural Networks**:
   - **Weight Decay**: Often implemented as L2 regularization.
   - **Dropout**: A technique where a fraction of neurons are randomly ignored during training, helping to prevent overfitting.
5. **Decision Trees and Ensemble Methods**:
   - **Random Forests**: Introduce randomness in feature selection and can include constraints like maximum tree depth.
   - **Gradient Boosting Machines (GBM)**: Can include regularization techniques in their loss functions, such as L1 or L2 penalties on leaf weights.
6. **Generalized Additive Models (GAMs)**:
   - Can incorporate penalties on the smoothness of the functions used to estimate relationships.
7. **K-Nearest Neighbors (KNN)**:
   - While KNN itself doesn't directly use regularization, techniques like feature selection and scaling can be considered forms of regularization in the broader sense.