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

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

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

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

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

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

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

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

10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.
A) True 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

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:-

1- a	2- a	3- b	4- b
5- c	6- b	7 - d	8 - d
9- a	10-b	11- a	12- a,b

MACHINE LEARNING

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

13. Explain the term regularization?

Answer:- Regularization is a technique used in machine learning to prevent overfitting by adding additional information to the model, typically in the form of a penalty against large coefficients. Overfitting occurs when a model learns the noise and details in the training data to an extent that it negatively impacts the model's performance on new data. Regularization helps to constrain or regularize the coefficient estimates towards zero to avoid this problem.

Types of Regularization:

1. L1 Regularization (Lasso Regression):

 - Adds the absolute value of the magnitude of the coefficients as a penalty term to the loss function.

 - Can result in sparse models where some coefficients are exactly zero, effectively performing feature selection.

 - The loss function for Lasso regression is:-

$$L = {RSS} + \lambda \sum_{j=1}^{p} | beta_j |$$

 where {RSS} is the residual sum of squares, \(\lambda \) is the regularization parameter, and beta_j are the coefficients.

2. L2 Regularization (Ridge Regression):-

 - Adds the square of the magnitude of the coefficients as a penalty term to the loss function.

 - Results in smaller, but non-zero, coefficients.

 - The loss function for Ridge regression is:-

$$L = {RSS} + \lambda \sum_{j=1}^{p} \beta_j^2$$

 where {RSS} is the residual sum of squares, (lambda) is the regularization parameter, and beta_j are the coefficients.

3. Elastic Net Regularization:-

- Combines both L1 and L2 regularization penalties.

- Useful when there are multiple correlated features.

- The loss function for Elastic Net is:

Purpose of Regularization:-

- Prevents Overfitting:- By adding a penalty to large coefficients, regularization prevents the model from becoming too complex and overfitting the training data.

- Improves Generalization:- A regularized model is more likely to perform well on new, unseen data.

- Feature Selection:- In the case of Lasso regression, it can automatically select relevant features by shrinking the coefficients of less important features to zero.

Regularization Parameter lambda:-

- The regularization parameter controls the strength of the penalty. A higher value of lambda increases the amount of shrinkage, leading to simpler models.

- It is typically selected using cross-validation to balance the trade-off between fitting the training data well and keeping the model coefficients small to ensure better generalization.

14. Which particular algorithms are used for regularization?

Answer:- Several algorithms and methods employ regularization techniques to improve model performance and prevent overfitting. Here are some key algorithms that utilize regularization:

1. Linear Regression with Regularization

- **Ridge Regression (L2 Regularization)**: Adds a penalty proportional to the square of the magnitude of the coefficients. This regularization helps in dealing with multicollinearity and prevents overfitting by shrinking the coefficients.
- **Lasso Regression (L1 Regularization)**: Adds a penalty proportional to the absolute value of the magnitude of the coefficients. It can lead to sparse models where some coefficients are exactly zero, thus performing feature selection.

- **Elastic Net Regression**: Combines both L1 and L2 regularization penalties, providing a balance between Ridge and Lasso regression. It is useful when there are many correlated features.

2. Logistic Regression with Regularization

- **Ridge Logistic Regression**: Similar to Ridge Regression but used for binary classification problems. It adds L2 regularization to the logistic loss function.
- **Lasso Logistic Regression**: Similar to Lasso Regression but used for binary classification problems. It adds L1 regularization to the logistic loss function.
- **Elastic Net Logistic Regression**: Combines both L1 and L2 regularization in the logistic loss function.

3. Support Vector Machines (SVM) with Regularization

- **Regularized SVM**: Adds a regularization term to the SVM objective function to control the trade-off between achieving a low error on the training data and minimizing the norm of the coefficients.

4. Generalized Linear Models (GLM) with Regularization

- **Regularized GLM**: GLM can include regularization terms like Ridge, Lasso, or Elastic Net in models such as Poisson regression, logistic regression, etc.

5. Neural Networks with Regularization

- **Dropout**: A technique where randomly selected neurons are dropped during training to prevent overfitting. This serves as a form of regularization for neural networks.
- **Weight Decay**: This is another form of L2 regularization applied to neural network weights to control the complexity of the model.
- **Early Stopping**: Monitoring the performance of the model on a validation set and stopping training when performance stops improving to prevent overfitting.

6. Regularized Decision Trees

- **Pruning**: While not regularization in the traditional sense, pruning decision trees can be seen as a regularization method that removes branches that have little importance.

Each of these algorithms incorporates regularization to enhance model performance and generalization by managing the complexity of the model and preventing overfitting.

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

Answer:- In linear regression, the term "error" refers to the difference between the actual observed values and the values predicted by the model. This concept is crucial for understanding how well the model is performing and how it can be improved. Here's a detailed breakdown:

Error in Linear Regression

1. **Definition**:
   - **Error** (or Residual): It is the difference between the observed value yiy_iyi and the predicted value
   - Here, yiy_iyi is the actual value, and y^i\hat{y}_iy^i is the predicted value from the linear regression model.
2. **Components of the Linear Regression Equation**: The linear regression model predicts the dependent variable yyy based on one or more independent variables XXX. The basic form of the linear regression equation is:
3. **Types of Errors**:
   - **Residual**: The error for an individual observation, as defined above.
   - **Total Error**: The sum of all residuals in the dataset, which can be measured using various metrics such as Sum of Squared Errors (SSE).
4. **Error Terms in the Model**:
   - The error term ϵi\epsilon_iϵi represents the difference between the true value of the dependent variable and the value predicted by the model. It captures the variability in the dependent variable that cannot be explained by the independent variables alone.
5. **Purpose of Minimizing Errors**:
   - **Least Squares Method**: The most common approach to fitting a linear regression model is to minimize the sum of the squared errors (residual sum of squares). This method ensures that the model parameters β0\beta_0β0 and β1\beta_1β1 are chosen such that the total squared difference between the observed and predicted values is minimized.
6. **Impact of Errors**:
   - **Model Fit**: The magnitude of the errors indicates how well the model fits the data. Large errors suggest that the model is not capturing the relationship well, while smaller errors suggest a better fit.
   - **Goodness of Fit**: Various metrics like R-squared, Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) are used to quantify the model's performance by assessing the errors.