MACHINE LEARNING

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

Ans. A) Least Square Error

2. Which of the following statement is true about outliers in linear regression?

Ans. A) Linear regression is sensitive to outliers

3. A line falls from left to right if a slope is ______?

Ans. B) Negative

4. Which of the following will have symmetric relation between dependent variable and independent variable?

Ans. B) Correlation

5. Which of the following is the reason for over fitting condition?

Ans. C) Low bias and high variance

6. If output involves label then that model is called as:

Ans. B) Predictive modal

7. Lasso and Ridge regression techniques belong to _________?

Ans. D) Regularization

8. To overcome with imbalance dataset which technique can be used?

Ans. D) SMOTE

9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

Ans. 9. A) TPR and FPR

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

Ans. False

11. Pick the feature extraction from below:

Ans. B) Apply PCA to project high dimensional data

12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

Ans. A) We don't have to choose the learning rate.

13. Explain the term regularization?

Ans. **Regularization** is a technique used in machine learning and statistics to prevent overfitting by adding a penalty to the loss function. The main goal of regularization is to improve the model's generalization to unseen data by discouraging overly complex models that may fit the training data too closely.

Key Concepts of Regularization:

1. **Overfitting**: This occurs when a model learns not only the underlying patterns in the training data but also the noise. Overfitting results in poor performance on new, unseen data.
2. **Penalty Terms**: Regularization introduces a penalty for large coefficients in the model. This discourages complexity:
   - **Lasso Regression (L1 Regularization)**: Adds the absolute value of the coefficients as a penalty term to the loss function. This can lead to sparse models, where some coefficients are exactly zero, effectively performing feature selection.
   - **Ridge Regression (L2 Regularization)**: Adds the square of the coefficients as a penalty term. It shrinks the

    coefficients but does not eliminate them entirely, leading to better prediction performance without feature selection.
  - **Elastic Net**: Combines both L1 and L2 regularization.

3. **Loss Function Modification**: The regularized loss function can be expressed as:

Regularized Loss=Original Loss +λ× Penalty\ text {Regularized Loss} = \ text {Original Loss} + \lambda \times \text {Penalty}Regularized Loss=Original Loss +λ× Penalty

where λ\ Lambda λ (the regularization parameter) controls the strength of the penalty. A larger value of λ\lambda λ imposes a greater penalty on the model complexity.

4. **Bias-Variance Trade off**: Regularization helps strike a balance between bias and variance. By penalizing complexity, it can increase bias slightly (since the model is constrained) but significantly decrease variance, leading to better overall performance

14. Which particular algorithms are used for regularization?

Ans. Regularization techniques can be applied to various machine learning algorithms to help prevent overfitting. Here are some of the key algorithms that utilize regularization:

1. Linear Regression

- **Lasso Regression (L1 Regularization)**: Adds a penalty equal to the absolute value of the magnitude of coefficients.
- **Ridge Regression (L2 Regularization)**: Adds a penalty equal to the square of the magnitude of coefficients.
- **Elastic Net**: Combines both L1 and L2 regularization.

2. Logistic Regression

- Similar to linear regression, logistic regression can also apply Lasso, Ridge, or Elastic Net regularization to improve its performance and reduce overfitting.

3. Support Vector Machines (SVM)

- SVM uses regularization to create a margin that separates the classes, helping to prevent overfitting, especially in high-dimensional spaces. The regularization parameter CCC controls the trade-off between maximizing the margin and minimizing the classification error.

4. Neural Networks

- **Weight Decay (L2 Regularization)**: A common technique where the loss function includes a penalty based on the weights of the network.
- **Dropout**: A form of regularization where randomly selected neurons are ignored during training, helping to prevent co-adaptation of hidden units.

5. Decision Trees and Ensemble Methods

- Regularization techniques can be applied in decision tree algorithms by limiting tree depth, the minimum samples required to split a node, and the minimum samples required at a leaf node.
- **Random Forests** and **Gradient Boosting** can also be regularized by controlling tree parameters and incorporating L1 or L2 penalties.

6. K-Nearest Neighbours (KNN)

- While KNN is not directly regularized, tuning the number of neighbours (K) can be seen as a regularization effect, as a larger K smoothens the decision boundary.

7. Principal Component Analysis (PCA)

- While PCA itself is primarily a dimensionality reduction technique, it can act as a form of regularization by reducing the number of features and focusing on the most significant components.

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

Ans. In the context of linear regression, **error** refers to the difference between the observed values (actual data points) and the values predicted by the regression model. This difference is crucial for understanding how well the model fits the data and is often analysed to improve the model's performance.

Key Concepts Related to Error in Linear Regression:

1. **Error Term**:
   - The error for a particular observation iii can be expressed mathematically as: Error I = y i − y^ I \text {Error}_i = yi - \hat{y}_i Error i=yi−y^i where:
     - Y I y i y i is the actual observed value.
     - y^ i\hat{y}_iy^i is the predicted value obtained from the linear regression equation.
2. **Linear Regression Equation**:
   - The basic form of a simple linear regression equation is: y^=β0+β1x+ϵ\hat{y} = \beta_0 + \beta_1 x + \epsilony ^ =β0+β1x+ϵ where:
     - y^\hat{y}y^ is the predicted value.
     - β0\beta_0β0 is the intercept.
     - β1\beta_1β1 is the coefficient for the independent variable xxx.
     - ϵ\epsilonϵ is the error term (the difference between the actual and predicted values).

3. **Types of Errors**:
   - **Residuals**: In the context of linear regression, the term "residual" refers to the error term. It is the difference between the actual value and the predicted value. Residuals are used to assess the fit of the model.
   - **Model Error**: This includes both bias and variance in the predictions made by the model. A model may have systematic errors (bias) or may be sensitive to variations in the training data (variance).
4. **Sum of Squared Errors (SSE)**:
   - A common metric to quantify the overall error of a regression model is the **Sum of Squared Errors**: SSE=∑i=1n(yi−y^i)2\text{SSE} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2SSE=i=1∑n(yi−y^i)2 where nnn is the number of observations. The SSE is minimized during the fitting process to achieve the best-fitting line.
5. **Mean Squared Error (MSE)**:
   - MSE is another common metric used to evaluate the average of the squared differences between the actual and predicted values: MSE=1n∑i=1n(yi−y^i)2\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2MSE=n1i=1∑n (yi−y^i)2
   - It provides a measure of the average error magnitude in the model's predictions.
6. **Root Mean Squared Error (RMSE)**:
   - RMSE is the square root of MSE, providing a measure of error in the same units as the response variable: RMSE=MSE\text{RMSE} = \sqrt{\text{MSE}}RMSE=M