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?

Answer: A) Least Square Error

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

Answer: A) Linear regression is sensitive to outliers

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

Answer: B) Negative

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

Answer: B) Correlation

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

Answer: C) Low bias and high variance

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

Answer: B) Predictive modal

7. Lasso and Ridge regression techniques belong to _____?

Answer: Regularization

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

Answer: D) SMOTE

9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary

classification problems. It uses _____ to make graph?

Answer: C) Sensitivity and Specificity

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

curve should be less.

Answer: A) True

11. Pick the feature extraction from below:

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?

Answer: 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.

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

13. Explain the term regularization?

Answer: One of the most common problems faced in Machine Learning is Overfitting. Regularization is used in machine learning as a solution to overfitting by reducing the variance of the model under consideration. Regularization can be implemented in multiple ways by either modifying the loss function, sampling method, or the training approach itself.

14. Which particular algorithms are used for regularization?

Answer: Algorithms used for regularization are Ridge Regression, LASSO (Least Absolute Shrinkage and Selection Operator) Regression and Elastic-Net Regression.

Ridge Regression: Ridge regression is a method for analyzing data that suffer from multi-collinearity.

LASSO Regression: LASSO is a regression analysis method that performs both feature selection and regularization in order to enhance the prediction accuracy of the model.

Elastic-Net Regression: Elastic-Net is a regularized regression method that linearly combines the L1 and L2 penalties of the LASSO and Ridge methods respectively.

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

Answer: Linear regression most often uses mean-square error (MSE) to calculate the error of the model. MSE is calculated by:

measuring the distance of the observed y-values from the predicted y-values at each value of x;

squaring each of these distances;

calculating the mean of each of the squared distances.

Linear regression fits a line to the data by finding the regression coefficient that results in the smallest MSE.