MACHINE LEARNING

In Q1 to Q11, only one option is correct, choose the correct option:

Q1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

Ans. Least Square Error

Q2. Which of the following statement is true about outliers in linear regression?

Ans. Linear regression is sensitive to outliers.

Q3. A line falls from left to right if a slope is _____?

Ans. Positive

Q4. Which of the following will have symmetric relation between dependent variable and independent variable?

Ans. Correlation

Q5. Which of the following is the reason for over fitting condition?

Ans. Low bias and high variance

Q6. If output involves label, then that model is called as:

Ans. Predictive model

Q7. Lasso and Ridge regression techniques belong to _____?

Ans. Regularization

Q8. To overcome with imbalance dataset which technique can be used?

Ans. SMOTE

Q9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?

Ans. TPR and FPR

Q10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

Ans. False

Q11. Pick the feature extraction from below:

Ans. Apply PCA to project high dimensional data.

In Q12, more than one options are correct, choose all the correct options:

Q12. 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.

B) It becomes slow when number of features is very large.

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

Q13. Explain the term regularization?

Ans. – Whenever our model gives less error in model evaluation and client feels that this model is overfitting their might have chances Train data and Test data is following a same pattern so, chances of getting a higher score are more in this situation the term regularization is used to check whether my model is overfitting or not.

Q14. Which particular algorithms are used for regularization?

Ans. Lasso (L1 form) and Ridge (L2 form) are the commonly used algorithms for regularization. Both are providing the best learning rate. It will help us reduce the overfitting problem. The main difference between Lasso and ridge is Lasso act as a feature selection tool where Ridge does not.

Q15. Explain the term error present in linear regression equation?

Ans. The term error in Linear Regression refers to Residual i.e., Actual – Predicted and there are situations when we have to display the overall error and for this, we have model evaluation techniques and these are-

Mean absolute error- In this you have to take every error and you take the average of same.

Mean squared error – In this you are going to take error square it and take mean of the Same and it ignore the outliers.it will go on majority.

Root mean squared error- It is the square root of mean squared error