

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?

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; A) Positive

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; C) Reinforcement learning

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; A) TPR and FPR

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

Ans; B) False

11. Pick the feature extraction from below:

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?

Ans; 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.

MACHINE LEARNING

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

13) Explain the term regularization?

Ans; In the Regularization technique, we reduce the magnitude of the independent variables by keeping the same number of variables. It maintains accuracy as well as a generalization of the model. It is a form of regression that shrinks the coefficient estimates towards zero. In other words, this technique forces us not to learn a more complex or flexible model, to avoid the problem of overfitting.

Mainly, there are two types of regularization techniques, which are given below:

1) Ridge Regression

2) Lasso Regression

14. Which particular algorithms are used for regularization?

Ans; Regularization is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting.

The commonly used regularization techniques are:

1) L1 regularization

2) L2 regularization

3) Dropout regularization

A regression model which uses L1 Regularization technique is called LASSO (Least Absolute Shrinkage and Selection Operator) regression.

A regression model that uses L2 regularization technique is called Ridge regression.

Lasso Regression adds “absolute value of magnitude” of coefficient as penalty term to the loss function(L). Ridge regression adds “squared magnitude” of coefficient as penalty term to the loss function(L)

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

Ans; In regression models, we assume that the relation between the response variable and predictors to be linear and we find out a straight line that fits the relation well.

The error term in regression is a catch-all for what we miss out with this model, because in reality

-The true relation may not be linear

-There may be other variables not included in the model that cause variation in response variable

-There may be measurement errors in the observations

The error is calculated as the difference between actual and estimated value of the response.

The error term of a regression equation represents all of the variation in the dependent variable *not* explained by the weighted independent variables.

A regression equation is the formula for a straight line — in this case, the best-fit line through a scatterplot of data. If there were no error, all the data points would be located *on* the regression line; to the extent they are not represents error; this is what the error term summarizes.