

MACHINE LEARNING

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

Answer – (a)

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

Answer – (b)

- 3) A-line falls from left to right if a slope is ____?

Answer –(b)

- 4) Which of the following will have symmetric relation between a dependent variable and independent variable?

Answer –(b)

- 5) Which of the following is the reason for overfitting condition?

Answer –(a)

- 6) if output involves label then that model is called as:

Answer –(a)

- 7) Lasso and Ridge regression techniques belong to _____?

Answer –(d)

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

Answer –(d)

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

Answer –(a)

- 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)

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

Answer –(b & c)

- 13) Explain the term regularization?

Answer – Regularization refers to techniques that are used to calibrate machine learning models in order to minimize the adjusted loss function and prevent overfitting or underfitting.

Using Regularization, we can fit our Machine learning model appropriately on a given test set and hence reduce the errors in it.

It is also used to reduce errors by fitting the function appropriately on the given training set and avoiding overfitting.

Sometimes the machine learning model performs well with the training data but does not perform well with test data. It means the model is not able to predict the output when dealing with unseen data by introducing noise in the output, and hence the model is called overfitting. This problem can be dealt with the help of a regularization technique.

The technique can be used in such a way that it will allow maintaining all variables or features in the model by reducing the magnitude of the variables. Hence, it maintains accuracy as well as a generalization of the model.

It mainly regularizes or reduces the coefficient of feature toward zero. In simple words, “In regularization technique, we reduce the magnitude of the feature by keeping the same number of features”

14) Which particular algorithms are used for regularization?

Answer – Understanding the use of regularization algorithms like LASSO, RIDGE and Elastic- Net regression.

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

Answer – An error term represents the margin of error within a statistical model: it refers to the sum of the deviation within the regression line, which provides an explanation for the difference between the theoretical value of the model and the actual observed results.

A regression line always has an error term because, in real life, independent variables are never perfect predictors of the dependent variables. Rather the line is an estimate based on the available data.

It is often said that the error term in a regression equation represents the effect of the variables that were omitted from the equation. This is unsatisfactory, even in simple contexts, as the following discussion should indicate.

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

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