MACHINE LEARNING

Answer 1
Which of the following methods do we use to find the best fit line for data in Linear Regression?
A) Lease square error

Answer 2
Which of the following statement is true about outliers in linear regression?
A) Linear regession is sensitive to outliers

Answer 3
A line falls from left to right if a slope is _____?
A) Positive

Answer 4
Which of the following will have symmetric relation between dependent variable and independent variable?
B) Correlation

Answer 5
Which of the following is the reason for over fitting condition?
C) Low bias and high variance

Answer 6
If output involves label then that model is called as
D) All of them

Answer 7
Lasso and Ridge regression techniques belong to _____?
D) regularization

Answer 8
To overcome with imbalance dataset which technique can be used?
D) SMOTE

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

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

Answer 11
Pick the feature extraction from below:
B) Apply PCA to project high dimensional data

Answer 12

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

A) we don't have to choose the learning rate
B) It becomes slow when number of features is very large

Answer 13

Explain the term regularization?

It is term used wrt to overfitting. Regularization is a technique to prevent the model from overfitting by adding extra information to it. Sometimes the machine learning model performs well with the training data but does not perform well with the test data. It means the model is not able to predict the output when deals with unseen data by introducing noise in the output, and hence the model is called overfitted. This problem can be deal with the help of a regularization technique.

This technique can be used in such a way that it will allow to maintain accuracy as well as a generalization of a model.

Answer 14

Which particular algorithms are used for regularization?

We use 3 algorithm for regularization

a. Ridge regression (L2 Norm))
b. Lasso (L1 Norm)
c. Dropout

Answer 15

Explain the term error present in linear regression equation?

In a linear regression i.e represented through y = mx + c . In a scatter plot, scatter represents data sets. we also have to find out a best fit line in the same scatter plot.

Residual is a vertical difference between the predicted y value and the overserved y value. We can also say that residual is the vertical difference between the data points and the best fit line.