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?

A) Least Square Error

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

A) Linear regression is sensitive to outliers

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

B) Negative

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

B) Correlation

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

A) High bias and high variance

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

B) Predictive modal

7. Lasso and Ridge regression techniques belong to _____?
D) Regularization

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

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

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

11. Pick the feature extraction from below:
D) Forward selection

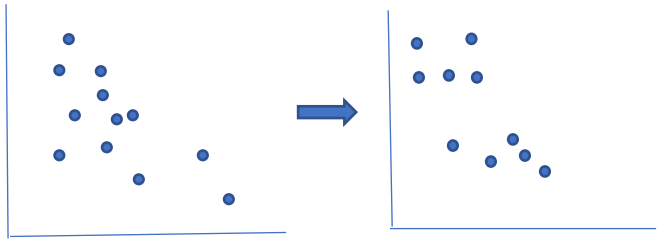
In Q12, more than one options are correct, choose all the correct options:
12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

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?

A: In machine learning we try to predict an output based on past trends of the data available to us. For this process, data given is categorised into two groups “predictors” and “predictand”. Predictand can also be called dependent variable as its value is associated on the values of the predictors or independent variables.
In instances, independent variables have large values but small in quantity making it complex and time consuming for machine to detect pattern, higher error function and calculate coefficients. For this purpose, we use REGULARIZATION as a technique to tune the independent variables by adding an additional penalty term in the error function. This additional penalty controls the excessive noise in the data and helps the machine to remove unnecessary complexity in functions. It helps us to achieve Occam’s razor (simplest form of model possible) or best fit model.



Before Regularization After Regularization

In all cases, our machine learning model follows principle of gradient descent. Depending on the hemisphere our present model lands of the graph describes it to be either under-fitting or over-fitting. In under-fitting the model fails the training phase and in over-fitting models become more rigid and excels the training phase but fails testing phase. Hence, regularization helps the model to achieve best fit.
Therefore, to conclude I would like to point out the importance of regularization. Regularization is mainly used to attain simplicity and accuracy of the machine learning model we are working with.

14. Which particular algorithms are used for regularization?

A: As we now know what regularization means, let us explore various techniques or algorithms used in this matter.
Types of regularization algorithms:

1st. Lasso Regularization or Least Absolute Shrinkage and Selection Operator (L1):
L1 adds penalty to the cost function(error). The penalty added is the sum of absolute value of weights. We can tune the parameters which decides how much to penalize the model. This regularization technique uses Manhattan distance which is computed by taking the sum of absolute value differences between the two vectors for defining the value of weight. Lasso helps in feature selection.

2nd. Ridge Regularization (L2):
Similar to L1, L2 too penalizes the cost function(error). The difference between the two is method of calculating the value of weights. This regularization technique uses Euclidean distance which is computed by taking the square root of the sum of squared differences between the two vectors for defining the value of weight. Ridge helps us to remove the problem of multicollinearity.

3rd. Elastic Net (L1.5):
This technique happens to be the combination of the above two to penalize the cost function(error). It uses both Euclidean and Manhattan distance to calculate the value of weight. Since elastic net is combination of L1 and L2 it gives us benefits of both the techniques but increases the cost.

Each technique has merits and their usage is entirely dependent on the data we are given.

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

A: Error in layman's term means the difference between the actual value and our predicted value. It is also called residual or cost function in some cases. Lesser the error better is our model.

For detailed explanation let us first understand linear regression equation which is given by:

Y= a + bx + e

Where,
Y- predicted value
a- intercept (mean value of Y when independent variable "x" is 0)
x- given value of an independent variable
b- slope (Rate of change in Y for unit change in x)
e- error term

Error term is very crucial in any linear regression model as it forms the basis for the accuracy and fit of the model.