1. C
2. C
3. C
4. A
5. A
6. B
7. A
8. ABCD
9. ABCD
10. ABD
11. Outlier is a term which refers to a point which stands out from the pattern of the points and can be distinguishable from the rest of the point which are there in the pattern.

Let us first understand the IQR , IQR stands for  **interquartile range** which is the difference between the 3rd quartile and 1st quartile .

As far as outliers detection using IQR is concerned , it is said that a datapoint is an outlier if

More than 1.5IQR above the third and below the first quartile .

Calculation: - Q3-1.5IQR =C , any number less than C will be categorised as outlier

Q3+1.5IQR=K, any number more than K will be categorised as an outlier.

----------------------------------------------

1. The primary difference between bagging and boosting is in bagging each model is built independently and the output is aggregated while in Boosting models are dependent on performance of previous model.

----------------------------------------------

1. Lets first understand R2

R2 =1-(SSres/SStot) where SSres is sum of square of errors.

Error = Actual – predicted

Where SStot is the sum of square of error when my predicted points is the the average of all the ouputs.Therefore error in this case will Predictedavg-Actual points.

R2 ranges from 0 to 1 , the more close to 1 means the line is best fitted.

Now in logistic regression when more and more features are added the value of R2 will be increasing always, therefore Adjusted R2 comes into picture to deal with this discrepancy and tries to penalise that are not correlated

R2 adjusted = (1-(1-R2) (N-1))/N-P-1

Where R2= sample R2 , P =number of independent features , N=total sample size.

Here suppose the features are increasing , menas P in increasing , which will minize the denominator and hence it will be able to decrease the adjusted R2 provide the features are not correlated.

-------------------------------------------------

1. Normalisation is a technique which scale the data under same format , for example if two features are having different different scale , Normalisation will bring them to the scale of 0 and 1.

Standardisation means converting all the dataspoint in such a way that the mean of the dataset becomes 0 and variance becomes 1.

-------------------------------------------------

1. Cross validation is a technique in which we train the model on all of our dataset , unlike train and split technique , Cross Validation is technique where for one time some parts of the dataset will act as train set and the rest of the dataset will act as test set , and this window of selecting the train and test set covers all the dataset periodically depending on the CV value.

Advantage – it becomes more accurate as whole dataset is covered and no information is ignored.

Disadvantage- The time complexity and space complexity becomes an important factor as CV technique has to rerun from scratch for K times and it impacts time and space.