PYTHON

1 (C) %

2(B) 0

3(C) 24

4(A) 2

5(D) 6

6(B)

7(A)

8(C)

9(A)(C)

10 (A)(B)

STATISTICS

1(A)

2(A)

3(D)

4(D)

5(C)

6(B)

7(B)

8(A)

9(C)

10 A normal distribution is a distribution which is in bell shape when plotting.Bell shape defines that the very low data is below the average and very low data is above the average . Mid-point which is the highest point represents thats most of the data is in average. At mid-point of the shape it represent that all parameters mean, median and mode all are equal here.

Normal distribution is symmetrical in shape.While moving left or right each value is 1 std away both left and right.

11 While treating with different type of datasets sometime some of the columns are empty or blank. And for good analysis we need to treat the missing values. For treating missing values we can use many methods. One basic method is fillna .With the help of fillna method we can fill missing values with mean ,median or mode data. Another method to treat missing values is imputer techniques.'SIMPLE IMPUTER' is the fast and approachable technique to fill missing values. We have to give strategy how we want to fill missing data either by mean, median or mode.

12 A/B testing a technique to comparing two version of website, app or advertising to find out which one is good from other.It is also known as split-testing.It helps to determine which one perform good and will give better results when optimizing. This technique is used by marketers and advertisers to find out which ad, email-campaigns is performing good in different segment of customers.

13 Mean imputation is technique used to treat missing values .But mean imputation is not widely accepted techniques.But according to me this technique is full based on dataset which we are dealing.For a good machine learning model we need good data for analysis but sometime mean imputation ignores or destroys the feature relationship with label. If we are dealing with dataset based on fitness of people during different age and between different sex groups. Mean analysis may fill missing value as mean of other data. Sometimes that data is irrelevant which may effect negatively on our model. But sometimes it also helps like, if we are dealing with height of people in a specific area. Height can be predicted or some little plus minus can be there. So according to my POV mean imputation is not that much acceptable technique but can be used.

14.Linear regression is one of the most fundamental algorithm and widely known machine learning algorithm which predict the regression analysis. It forms a relationship between multiple features to the label using a straight best fit line(regression line).

15. Statistic is having majorly 2 branches.

i) Descriptive statistic- In this type of statistic we deal with the collection, orgnaizing, and plotting of sample data with the help of the tables, graphs and paragraphs.

ii) Inferential statistics – inferential statistics involves the conclusion of analysis done with the study of data.

MACHINE LEARNING

1(A)

2(A)

3(B)

4(B)

5(C)

6(B)

7(D)

8(D)

9(A)

10(B)

11(B)

12(A)(B)(C)

13 Regularization is a term used to reduce the complexity of the model and help our model to avoid overfitting .By using regularization techniques we give some terms to the model and try to sort out complex features.

14 LASSO(L1) and RIDGE(L2) are 2 most popular algorithms in regularization to avoid the complexity of the model. L1 or LASSO try of avoid the features which are not or having less relationship with label while L2 or RIDGE try to give a little preference to the feature according to their relationship level.

15 Error is a term used to represent the margin of expected and predicted value. The higher the margin the higher the error. In Linear regression we used RMSE & MSE technique to find out how much error our model is getting. RMSE

is root mean square error and MSE is mean squared error. MSE uses the distamce between the observed y-value to the predicted y-value per value of x. RMSE is root of MSE.