10

DATA SCIENCE

INTERVIEW QUESTIONS

ASKED BY

Google

#1

What will you do if removing missing values from a dataset cause bias?

comment your answer

#2

Cars are implanted with speed tracker so that the insurance companies can track about our driving state. Based on this new scheme what kind of business questions can be answered?

#3

You have 10 coins. You toss each coin 10 times (100 tosses in total) and observe results. Would you modify your approach to the the way you test the fairness of coins?

#4

Why use feature selection? If two predictors are highly correlated, what is the effect on the coefficients in the logistic regression? What are the confidence intervals of the coefficients?

#5

When using Gaussian mixture model, how do you know it is applicable? (Normal distribution)

#6

If you could get the dataset on any topic of interest, irrespective of the collection methods or resources then how would the dataset look like and what will you do with it?

#7

If each of the two coefficient estimates in a regression model is statistically significant, do you expect the test of both together is still significant?

#8

You have a google app and you make a change. How do you test if a metric has increased or not?

#9

If the labels are known in the clustering project, how to evaluate the performance of the model?

#10

What is the difference between a bagged model and a boosted model?