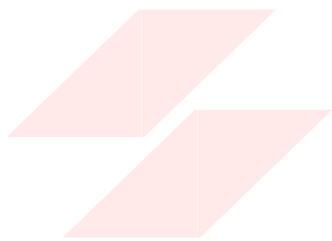


9. Suppose we have a dataset which have two classes A and B. The percentage of class A is 40% and percentage of class B is 60%. Calculate the Gini index and entropy of the dataset.
10. What are the advantages of Random Forests over Decision Tree?
11. What is the need of scaling all numerical features in a dataset? Name any two techniques used for scaling.

12. Write down some advantages which scaling provides in optimization using gradient descent algorithm.
13. In case of a highly imbalanced dataset for a classification problem, is accuracy a good metric to measure the performance of the model. If not, why?
14. What is “f-score” metric? Write its mathematical formula.
15. What is the difference between `fit()`, `transform()` and `fit_transform()`?



# FLIP ROBO