Friday, March 22, 2024 15:51

1. Suppose we have a data set with k features and N samples and N ≫ k then:

   a. In the event of discrete features (2 category only), what is the highest number of leaves in decision tree?
   b. In the event of continuous features, what is the highest number of leaves in decision tree?

[a] 2^k,

In the event of discrete features, each feature has only 2 categories. Since each feature can split the data into two based on categories, for k features there could be upto 2^k leaves in decision tree.

[b] N,

with N samples, assuming that every sample has a unique value for each continuous feature, the maximum number of leaves could be as high as N.