DECISION TREES:

* If we go for data which we make, like if-else statements, and classify the data, then we use decision trees.
* For example, to play or stay at home, or go outside, depending on the type of weather
* Geometric intuition of trees:
* Plot the graph and divide by using the line or plane then make a mathematical model, and it can be done recursively.
* The first picking of the feature, such that it will be the best.

Terminology:

* Root node
* Splitting
* Branch/subtree
* Decision node
* Leaf node

Entropy: The misuse of disorder or purity/impurity.

Entropy(S)=−(∑​pi​log2​(pi​)(i=1 to n)

N = no of class labels

* The greater the entropy, the greater is uncertainty
* We can use base e or 2 for log
* The graph for entropy and probability is the reverse bowl.

Information gain:

Information Gain = Entropy (before split)−Entropy (after split)

* It uses a recursive greedy search algorithm and splits it into groups
* If the entropy for the group is zero we touch or modify that group

Gini Impurity:

Gini(S)=1−(i=1 to n)∑​pi^2​

The maximum value of the Gini is 0.5