**Decision Tree Algorithm Implementation**

Node is a feature column.

Leaf is the Label value.

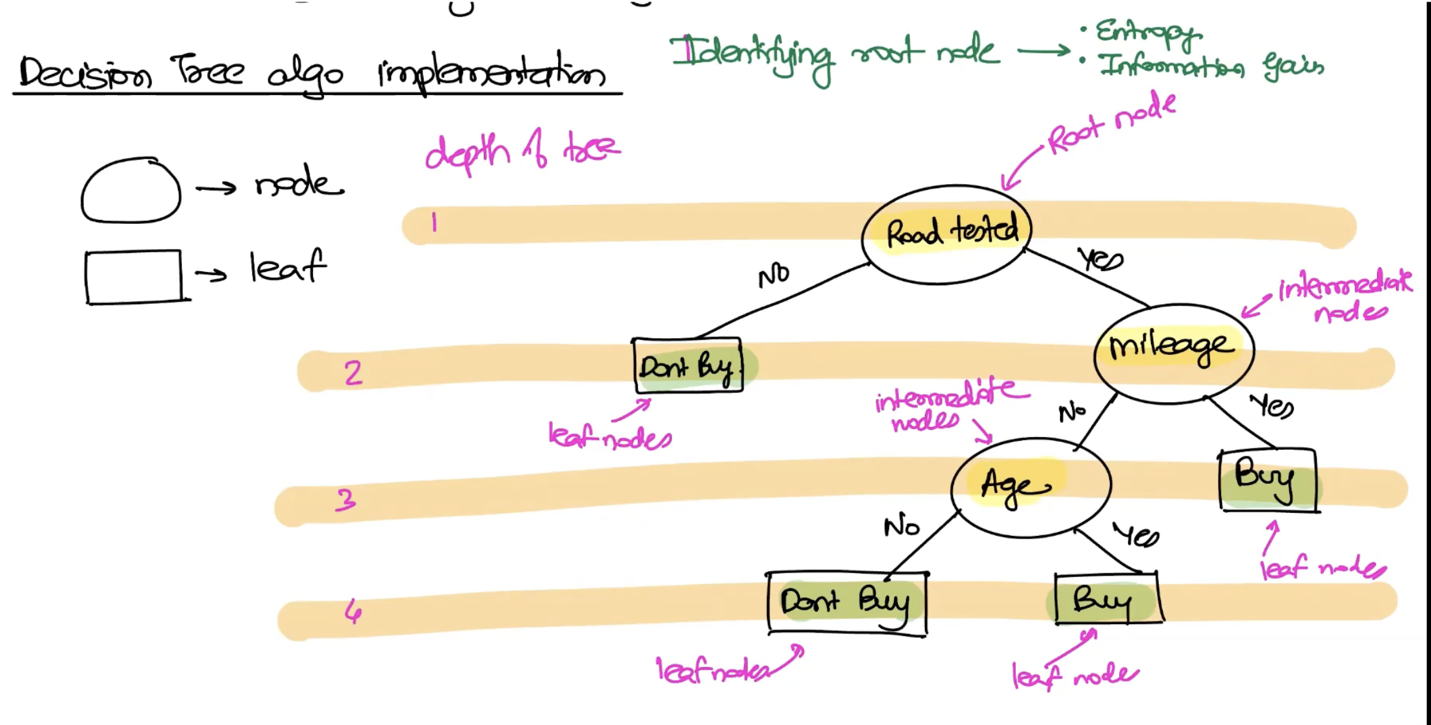
When we talk about Machine Learning algorithm with Decision Tree, the ultimate goal of training is, identify the root node and of course create decision tree with certain depth.

For identify the root node, there are two formulas: (criterion parameters)

* Entropy formula 🡺 entropy
* Information gain formula 🡺 gini

There are other Parameter in using Decision Tree: max\_depth.

max\_depth is very important when we are dealing with overfitting. Minimum value for max\_depth is 2.



When you train your model:

* first train it in a default way and see you get overfitted?

If you are getting overfitted:

* then play with max\_depth parameter: increase the max\_depth(start with 2 and increase it) to see you don’t have any overfitted issue.