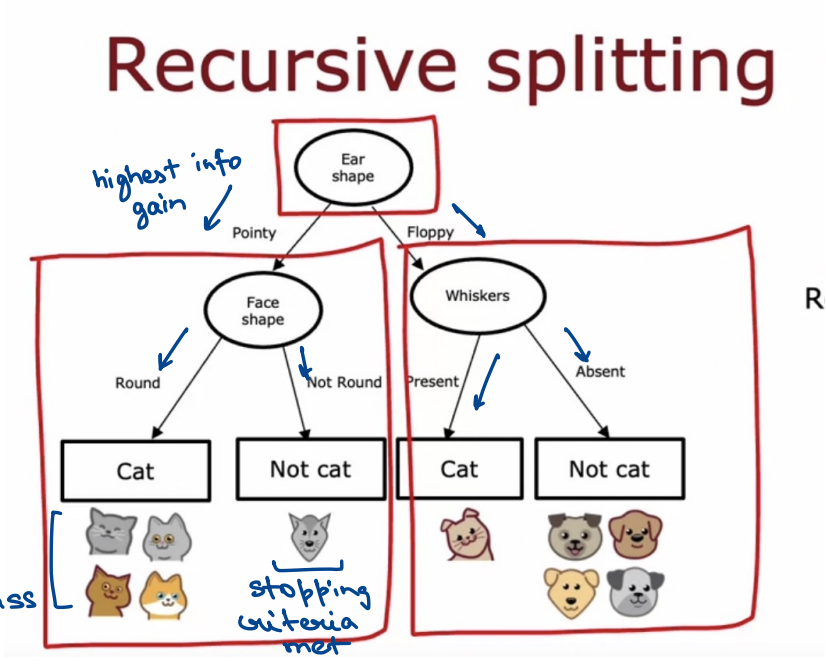
This is a basic layout of how a decision tree is created :-

- Start with all the training examples at the root node.

- Calc. information gain to check which feature gives you the highest information gain and then choose it.

- Split the tree left and right branches with the no. of training examples at each branch depending on the feature chosen.

- Keep repeating until a stopping criteria is met :
  - → A node is 100% one class (i.e. entropy is 0)
  - → Tree has reached the maximum depth which was set
  - → Information gain from additional splits is less than threshold.
  - → No. of examples in a node is less than threshold.



R

It is basically a recursion algorithm (calls itself). There are open-source libraries that make it easier to choose a maximum depth parameter

You could use a cv set for choosing a maximum depth parameter because it is kinda like having a higher degree polynomial, but open source libraries are better.