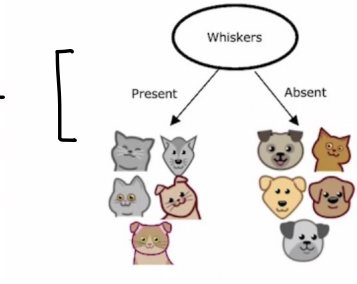
Using a single decision tree is prone to small changes in data.

Trees are highly sensitivity to small changes of the data

Suppose we have this tree

if we change only this one cat all of the subsets will change

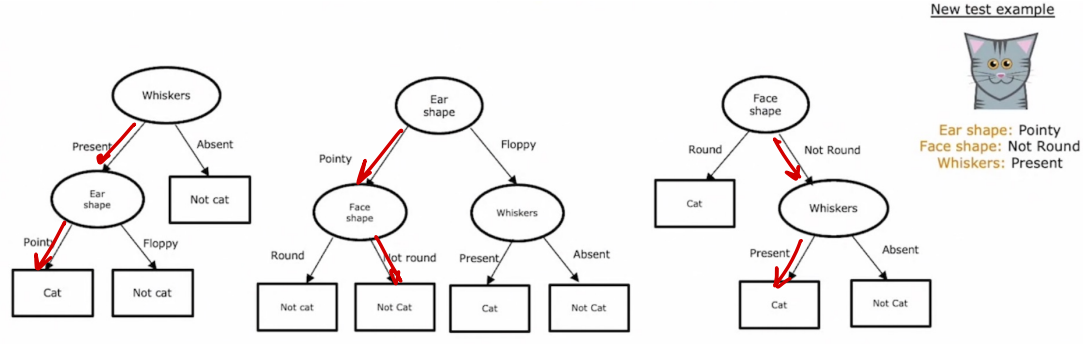
instead of having pointy ears, round face and whiskers absent it has floppy ears, round face and whiskers present.



totally different split ⇒ totally different trees.

To avoid this we train a bunch of decision trees and take a vote.

We have what is called a tree ensemble.



Prediction : cat Prediction : Not cat Prediction : cat

Basically, is a collection of trees and we'll look at the majority of votes.

Since majority votes were for cat, therefore we know the test example.