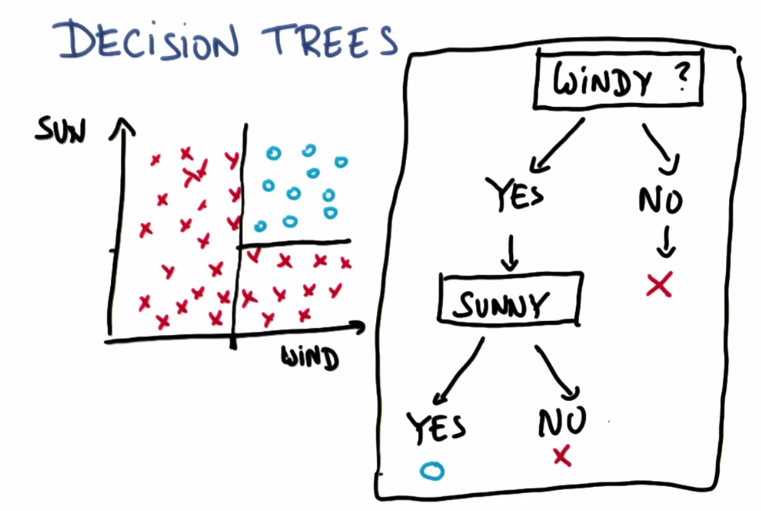
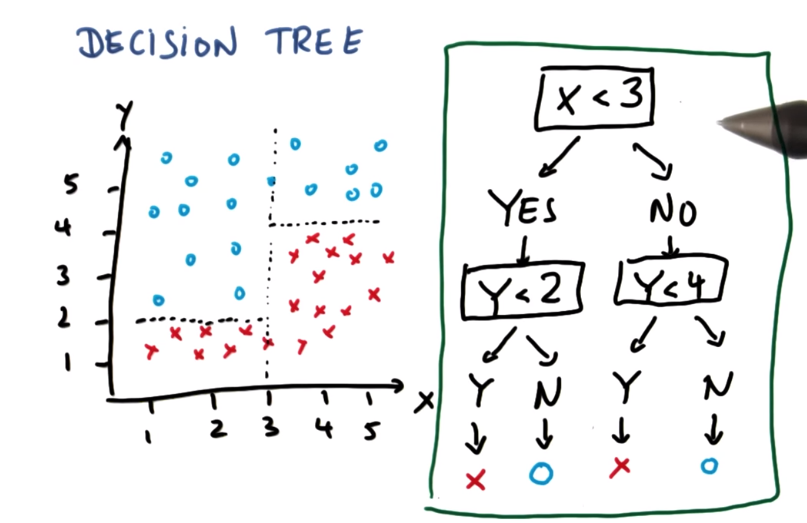
Decision Trees:

Let you make non-linear decision making with simple linear decision surfaces.

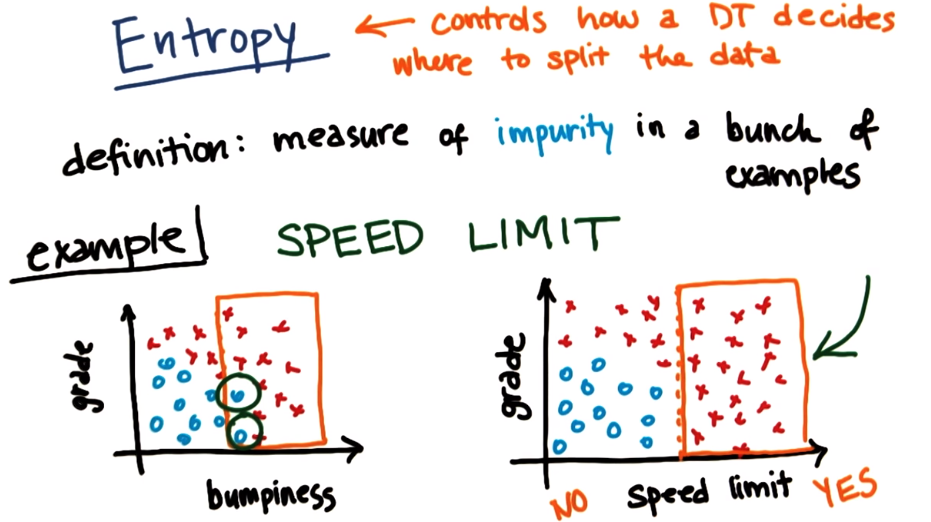
They do this by asking multiple linear questions.

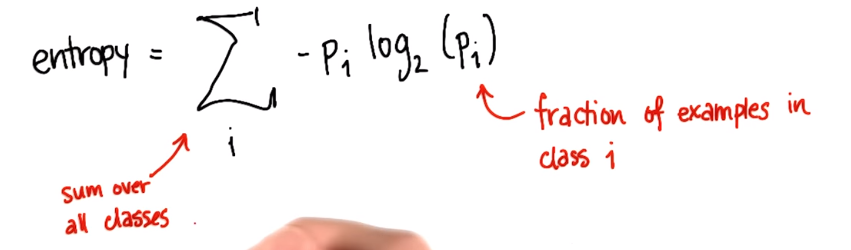




The algorithm defines these boundaries on its own.

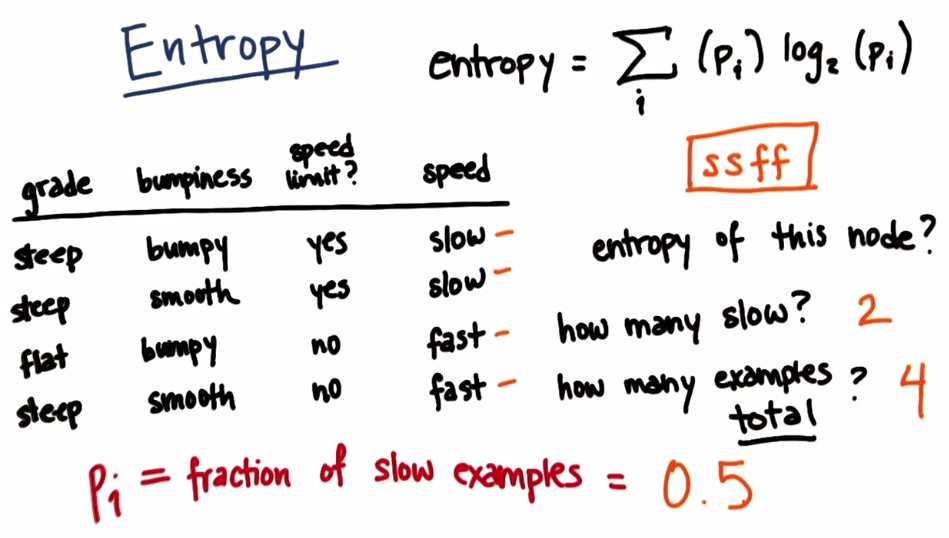
Decision Tree Entropy

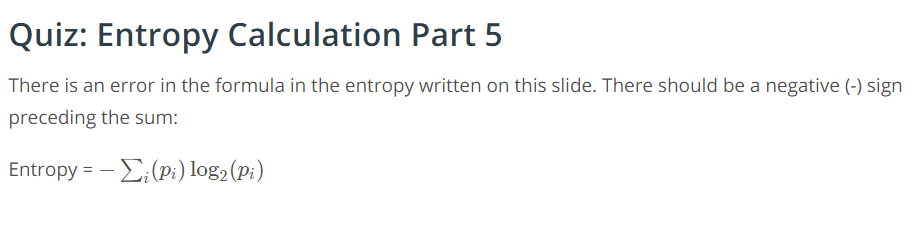


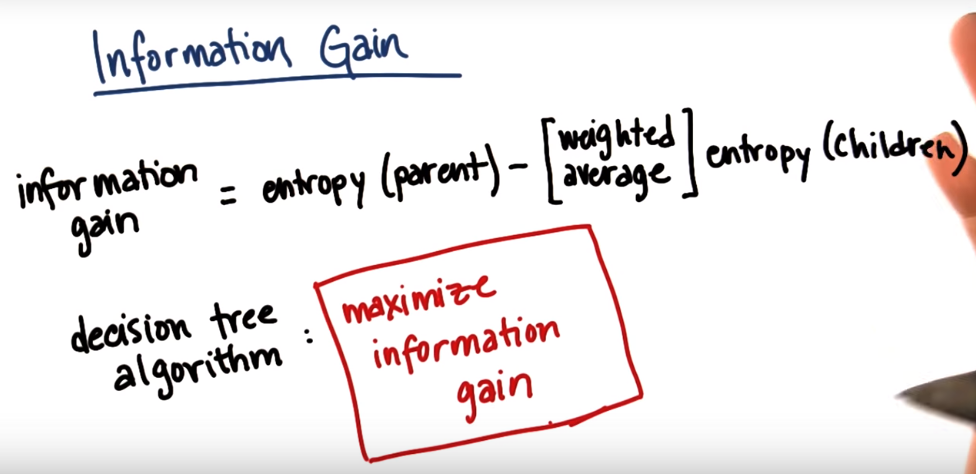


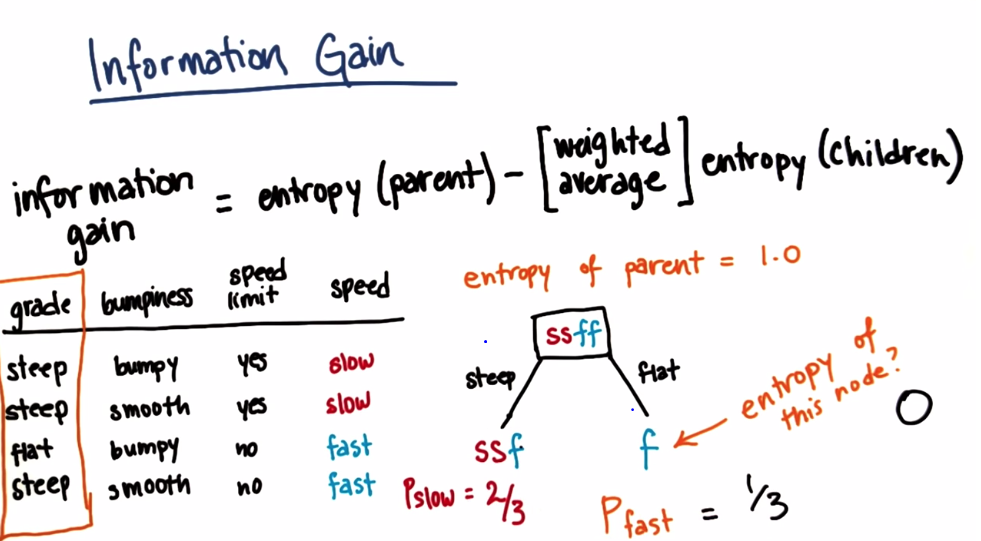
Entropy is zero if all the classes are the same, in a set, and 1 if they are all not.

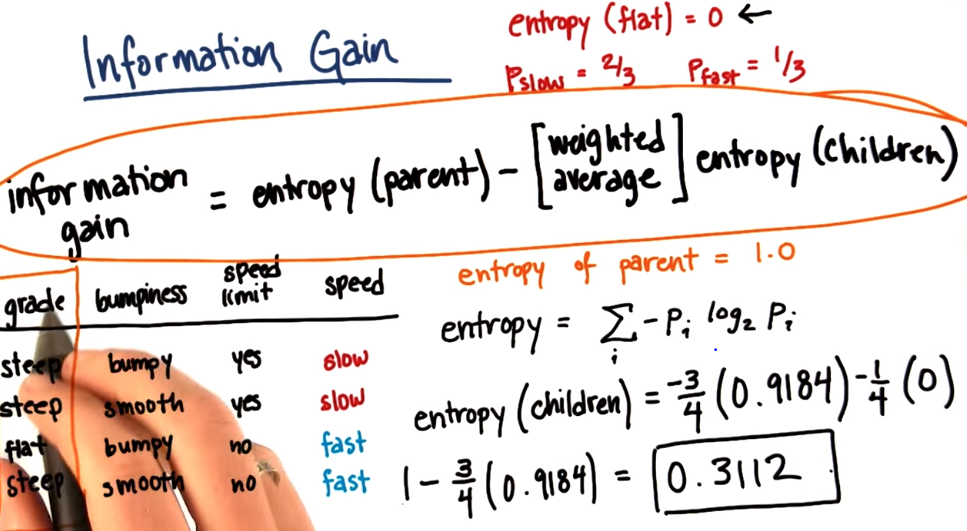
Entropy









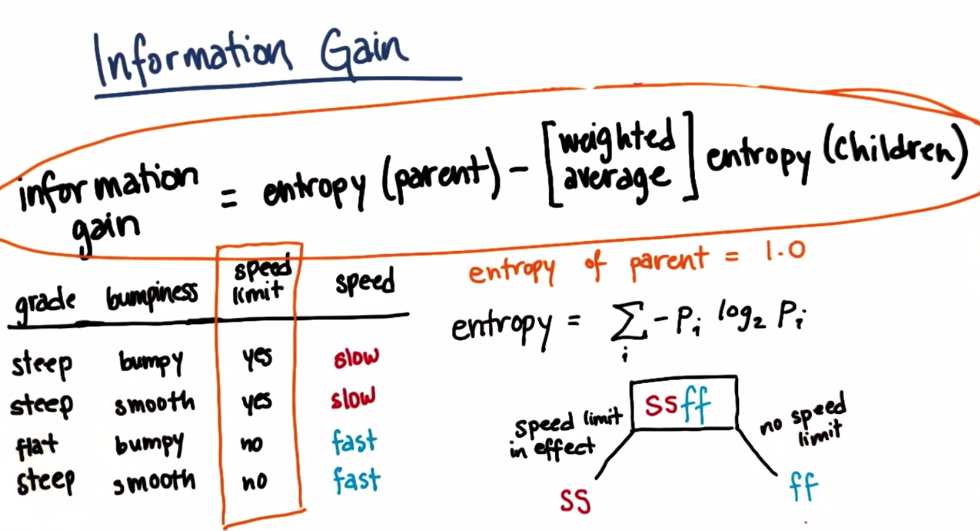


Information gain as a result of splitting based on the grade.

We can also split based on the bumpiness: which gives an information gain of 0.

Therefore this training set is not split by bumpiness.

Information gain from speed limit is 1. Therefore speed limit is what the decision tree splits on.



Tuning parameters:, in the decision tree classifier, the default criterion is called gini, but it can also take entropy.

A bias in a machine learning algorithm there is a dilemma called the bias variance dilemma.

If the algorithm is too biased, it will not react to the data it has seen before which is very bad, but if the bias is very low, then the ability for your algorithm to react to something it hasn’t seen before is very bad. This means that we will want to aim for something in the middle for a great machine learning algorithm. You want your algorithm to have the right bias to generalize but still open to listen to the data it learned.

Strengths:

Easy to implement.

Weakness:

Prone to over fitting, especially if there are too many features. (Stop at the appropriate time)