

A Decision Tree Learning Example

Example	Attribute			Class
	<i>Color</i>	<i>Shape</i>	<i>Size</i>	
1	red	square	big	+
2	blue	square	big	+
3	red	round	small	−
4	green	square	small	−
5	red	round	big	+
6	green	round	big	−

Initial Call: DTL([1, 2, 3, 4, 5, 6], [Color, Shape, Size], +)

Finding the Best Attribute

$$\text{Remainder}(\text{Color}) = \frac{3}{6}I\left(\frac{2}{3}, \frac{1}{3}\right) + \frac{1}{6}I\left(\frac{1}{1}, \frac{0}{1}\right) + \frac{2}{6}I\left(\frac{0}{2}, \frac{2}{2}\right) = 0.459$$

$$\text{Gain}(\text{Color}) = I\left(\frac{3}{6}, \frac{3}{6}\right) - \text{Remainder}(\text{Color}) = 1 - 0.459 = 0.541$$

$$\text{Remainder}(\text{Shape}) = \frac{3}{6}I\left(\frac{2}{3}, \frac{1}{3}\right) + \frac{3}{6}I\left(\frac{1}{3}, \frac{2}{3}\right) = 0.918$$

$$\text{Gain}(\text{Shape}) = I\left(\frac{3}{6}, \frac{3}{6}\right) - \text{Remainder}(\text{Shape}) = 1 - 0.918 = 0.082$$

$$\text{Remainder}(\text{Size}) = \frac{2}{3}I\left(\frac{3}{4}, \frac{1}{4}\right) + \frac{1}{3}I\left(\frac{0}{2}, \frac{2}{2}\right) = 0.541$$

$$\text{Gain}(\text{Size}) = I\left(\frac{3}{6}, \frac{3}{6}\right) - \text{Remainder}(\text{Size}) = 1 - 0.541 = 0.459$$

Hence, Color is best because it maximizes the Gain (or, equivalently, minimizes the Remainder)

So, recurse on

DTL([2], [Shape, Size], +)
 DTL([4, 6], [Shape, Size], −)
 and DTL([1, 3, 5], [Shape, Size], +)