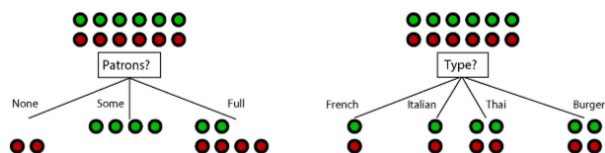


# Part2

Wednesday, 20 January 2021 10:26 am

Activity 1

## Choosing an Attribute



$$\begin{aligned} \text{For Patrons, Entropy} &= \frac{1}{6}(0) + \frac{1}{3}(0) + \frac{1}{2} \left[ -\frac{1}{3} \log\left(\frac{1}{3}\right) - \frac{2}{3} \log\left(\frac{2}{3}\right) \right] \\ &= 0 + 0 + \frac{1}{2} \left[ \frac{1}{3}(1.585) + \frac{2}{3}(0.585) \right] = 0.459 \end{aligned}$$

$$\text{For Type, Entropy} = \frac{1}{6}(1) + \frac{1}{6}(1) + \frac{1}{3}(1) + \frac{1}{3}(1) = 1$$

(Will cover if there is some time left after covering the assignment part)

Consider the task of predicting whether children are likely to be hired to play members of the Von Trapp Family in a production of The Sound of Music, based on these data:

height	hair	eyes	hired
short	blond	blue	+
tall	red	blue	+
tall	blond	blue	+
tall	blond	brown	-
short	dark	blue	-
tall	dark	blue	-
tall	dark	brown	-
short	blond	brown	-

1.

Compute the information (entropy) gain for each of the three attributes (height, hair, eyes) in terms of classifying objects as belonging to the class, + or -.

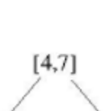
2.

Construct a decision tree based on the minimum entropy principle.

## Laplace Pruning

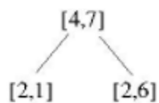
$$E = 1 - \frac{n+1}{N+k}$$

*N is the total # of items in current node*  
*n is the # of items in the majority class*  
*k is # of classes*



Error of Parent

$$= 1 - \frac{7+1}{13} = 1 - \frac{8}{13} = \frac{5}{13} \approx 0.385$$



Error of Parent

$$= 1 - \frac{7+1}{1+2} = 1 - \frac{8}{13} = \frac{5}{13} \doteq 0.385$$

Error (left child)

$$= 1 - \frac{2+1}{3+2} = 1 - \frac{3}{5} = \frac{2}{5} = 0.4$$

Error (right child)

$$= 1 - \frac{6+1}{8+2} = 1 - \frac{7}{10} = \frac{3}{10} = 0.3$$

Error (children expected):

$$= \frac{3}{11} \times 0.4 + \frac{8}{11} \times 0.3 \doteq 0.327 < 0.385$$

$\therefore$  don't prune.

Activity 3:

• self-exercise:

\* exactly the same process but only more features

\* ask me via email or post on the forum if you're stuck, also check the solution which Anne will post later.