Decision Tree Learning

1. The next attribute chosen by ID3 will be the Est attribute, the estimated wait time quoted.
2. The expected information gain associated with the Est attribute came out to 0.4286 (the next highest was 0.3207)
3. I calculated the expected information gain as B(T/(T+F)) - Remainder(Est), where Remainder(Est) = Σ(Tk+Fk)/(T+F)*B(Tk/(Tk+Fk), Tk and Fk being the number of T and F for each possible value of Est, and summed over all possible values of Est, and B(q) = -(q$\log_2$q + (1-q)$\log_2$(1-q)). For Est, the Tk and Fk breakdown was 3:0 for 0-10min, and 1:1 for 10-30min and 30-60min, and we had no examples with Est=>60 and Hun=T

Trace Backpropagation

1.
    a. $in_i = \Sigma w_{h,i} a_h$ over all connected inputs a
    b. $a_i = g(in_i)$ or after a substitution with the previous equation, $a_i = g(\Sigma w_{h,i} a_h)$
    c. $\Delta[i] = g'(in_i) \Sigma w_{i,j} \Delta[j]$ for all nodes j in the next layer in the network
    d. $W_{i,j} = w_{i,j} + \alpha \times a_i \times \Delta[j]$
2.

Iteration	a_c	$\Delta[c]$	a_d	$\Delta[d]$	w_{0c}	w_{ac}	w_{bc}	w_{cd}	w_{0d}
x1	0.5987	0.0049	0.5793	0.1025	0.3	0.2015	0.2000	0.2184	0.3
x2	0.5987	-0.0074	0.5819	-0.1416	0.3	0.2015	0.1978	0.1930	0.3