

Decision Tree Hw

Paul Johnston

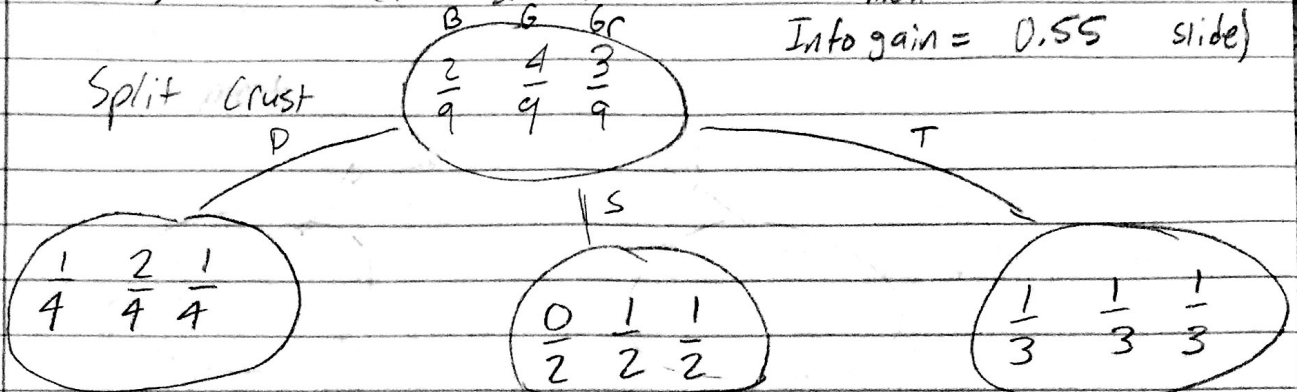
CS 478

2/19/2019

$Inf(s) = 1.53$ (from slide)

$Info_{meat}(s) = 0.98$ (from slide)
 $Info_{gain} = 0.55$

Split Crust

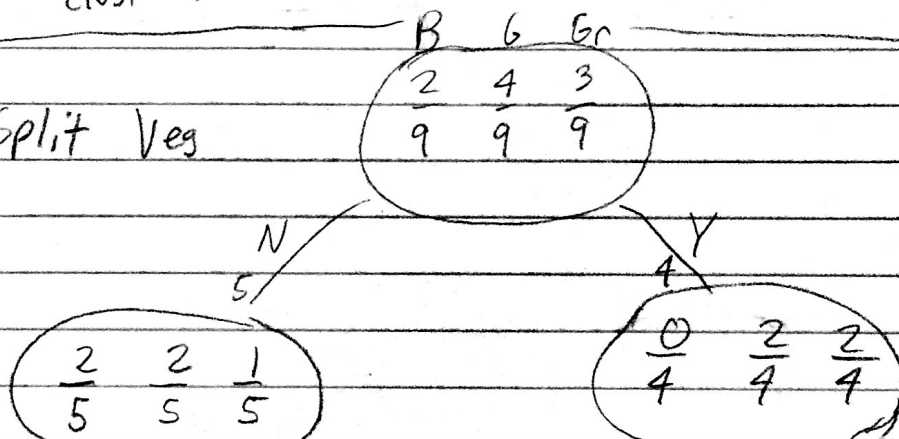


$$\frac{4}{9} \left(-\frac{1}{4} \log_2 \left(\frac{1}{4} \right) - \frac{2}{4} \log_2 \left(\frac{2}{4} \right) - \frac{1}{4} \log_2 \left(\frac{1}{4} \right) \right) + \frac{2}{9} \left(-\frac{1}{2} \log_2 \left(\frac{1}{2} \right) - \frac{1}{2} \log_2 \left(\frac{1}{2} \right) \right) + \frac{3}{9} \left(-\frac{1}{3} \log_2 \left(\frac{1}{3} \right) - \frac{1}{3} \log_2 \left(\frac{1}{3} \right) - \frac{1}{3} \log_2 \left(\frac{1}{3} \right) \right)$$

$$0.667 + 0.22 + 0.53$$

$$Info_{crust}(s) = 1.42 \quad Information\ Gain = 0.11$$

Split Veg



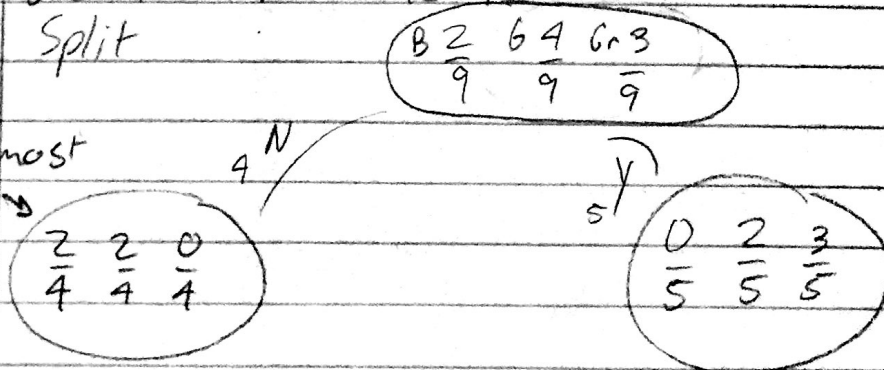
$$\frac{5}{9} \left(-\frac{2}{5} \log_2 \left(\frac{2}{5} \right) - \frac{2}{5} \log_2 \left(\frac{2}{5} \right) - \frac{1}{5} \log_2 \left(\frac{1}{5} \right) \right) + \frac{4}{9} \left(-\frac{2}{4} \log_2 \left(\frac{2}{4} \right) - \frac{2}{4} \log_2 \left(\frac{2}{4} \right) \right)$$

$$0.85 + 0.44$$

$$Info_{veg}(s) = 1.29 \quad Information\ gain = 0.24$$

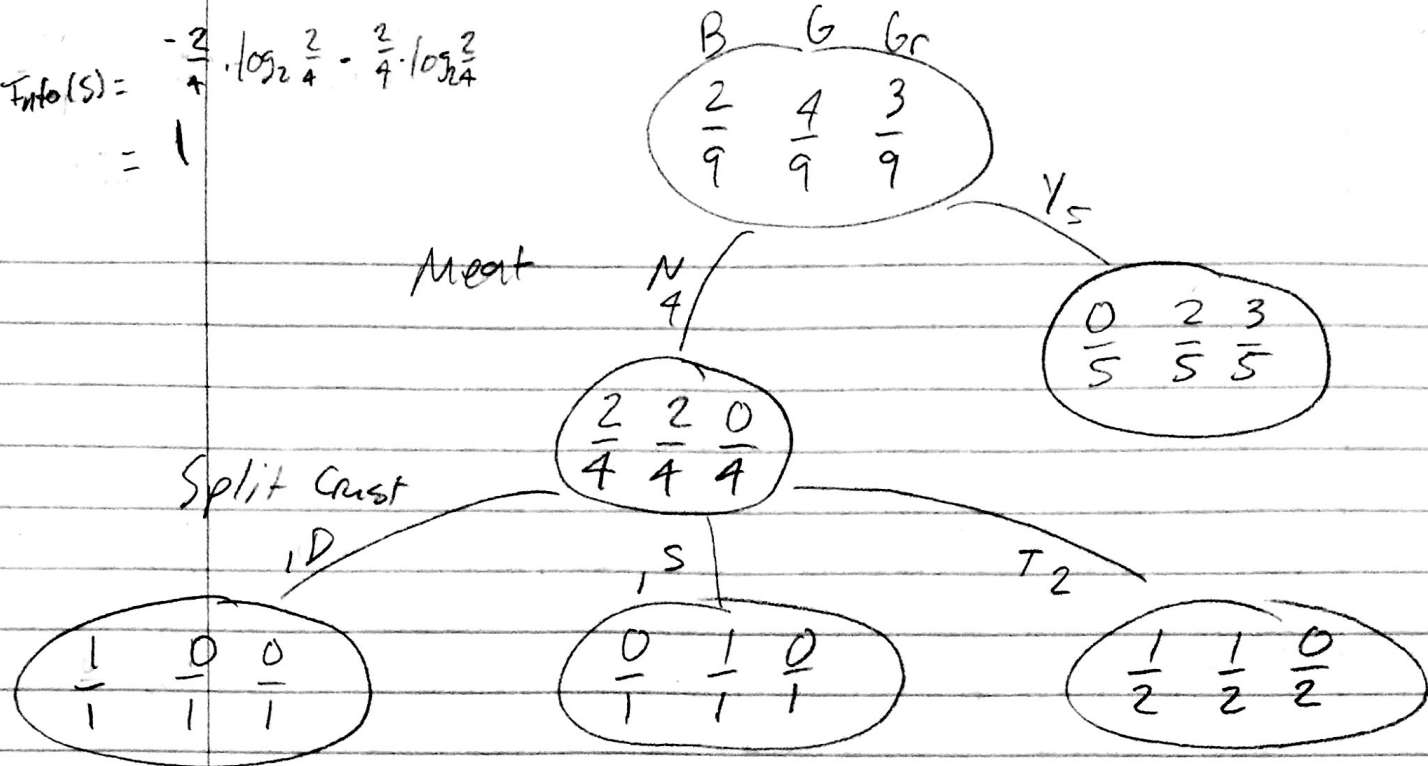
Best attribute: Meat

Split



Left most node

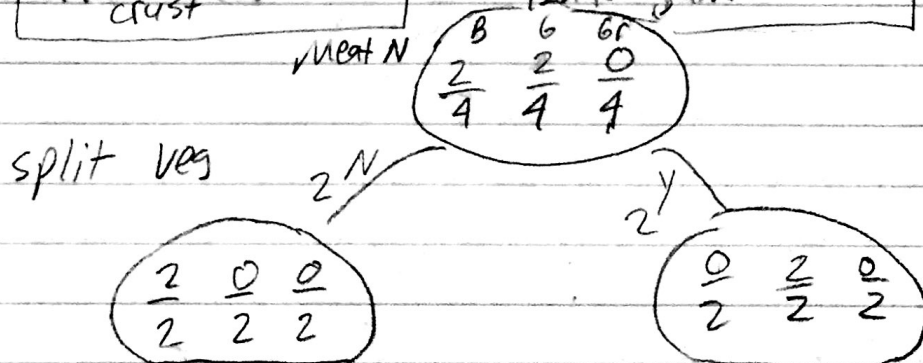
$$Info(S) = -\frac{2}{4} \cdot \log_2 \frac{2}{4} - \frac{2}{4} \cdot \log_2 \frac{2}{4} = 1$$



$$\frac{1}{4} \cdot \left(-\frac{1}{1} \cdot \log_2 \left(\frac{1}{1} \right) \right) + \frac{1}{4} \cdot \left(-\frac{1}{1} \cdot \log_2 \left(\frac{1}{1} \right) \right) + \frac{2}{4} \cdot \left(-\frac{1}{2} \cdot \log_2 \left(\frac{1}{2} \right) - \frac{1}{2} \cdot \log_2 \left(\frac{1}{2} \right) \right)$$

0 + 0 = 0.5

$Info_{crust}(S) = 0.5$ $Info_{gain} = 0.5$



$$\frac{2}{4} \cdot \left(-\frac{2}{2} \cdot \log_2 \left(\frac{2}{2} \right) \right) + \frac{2}{4} \cdot \left(-\frac{2}{2} \cdot \log_2 \left(\frac{2}{2} \right) \right)$$

$Info_{veg}(S) = 0$

Information gain = 1

Best attribute: Veg