C4.5

$$\text{Gain ratio} = \frac{\text{Gain}(A)}{\text{SplitInfo}(A)}$$

$$\text{SplitInfo}(A) = \sum_{i=1}^{v} \frac{|D_j|}{|D|} * \log_2 \frac{|D_j|}{|D|}$$

$$\text{Split Info}(\text{Age}) = -\frac{5}{14}\log_2\frac{5}{14} - \frac{4}{14}\log_2\frac{4}{14} - \frac{5}{14}\log_2\frac{5}{14}$$

$$= 1.57$$

$$\text{Gain ratio}_{(\text{age})} = \frac{\text{Gain}(\text{Age})}{\text{SplitInfo}(\text{Age})} = \frac{0.246}{1.57} = 0.156$$

$$\text{SplitInfo}(\text{Income}) = -\frac{4}{14}\log_2\frac{4}{14} - \frac{6}{14}\log_2\frac{6}{14} - \frac{4}{14}\log_2\frac{4}{14} = 1.75$$

$$\text{Gain Ratio}_{(\text{income})} = \frac{\text{Gain}(\text{income})}{\text{SplitInfo}(\text{income})} = \frac{0.035}{1.75} = 0.018$$

Page: Date: / /

$$\text{splitInfo}(\text{student}) = -\frac{7}{14}\log_2\frac{7}{14} - \frac{7}{14}\log_2\frac{7}{14} = 1$$

$$\text{Gain Ratio}_{(\text{student})} = \frac{\text{Gain student}}{\text{splitInfo(student)}} = \frac{0.16}{1} = 0.16$$

$$\text{split Info}(\text{credit-rating}) = -\frac{8}{14}\log_2\frac{8}{14} - \frac{6}{14}\log_2\frac{6}{14} = 0.98$$

$$\text{Gain Ratio}_{(c\text{-}r)} = \frac{\text{Gain c-r}}{\text{splitInfo(c-r)}} = \frac{0.05}{0.98} = 0.05$$