

Quiz 9: info 3≡3

| R |
|----|
| 6 |
| 7 |
| 8 |
| 9 |
| 13 |

No
Yes
No
Yes
Yes

| | | | | |
|---|---|---|---|----|
| N | Y | N | Y | Y |
| 6 | 7 | 8 | 9 | 13 |

$$Info_{age}(D) = \frac{5}{14} I(2,3) + \frac{4}{14} I(4,0) + \frac{5}{14} I(3,2) = 0.694$$

$$Info(D) = - \sum_{i=1}^m p_i \log_2(p_i)$$

needed (after using A to split D into v partitions) to classify D:

$$Info_A(D) = \sum_{j=1}^v \frac{|D_j|}{|D|} \times Info(D_j)$$

gained by branching on attribute A

$$Gain(A) = Info(D) - Info_A(D)$$

Sol. $Info(D) = -\frac{1}{9} I(9,1) + \frac{2}{9} I(8,2) + Info_A(D) 2$