HW-2

Pavan Santosh

March 2024

1 Problem 1

1.1

Measuring conditional entropy on each of X1 attributes
For X1=1 branch

$$-(\frac{2}{2+1+1}\log_2(\frac{2}{2+1+1})+\frac{1}{2+1+1}\log_2(\frac{1}{2+1+1})+\frac{1}{2+1+1}\log_2(\frac{1}{2+1+1}))$$

1.5

For X1=0 branch

$$-(\frac{0}{0+1+1}\log_2(\frac{0}{0+1+1})+\frac{1}{0+1+1}\log_2(\frac{1}{0+1+1})+\frac{1}{0+1+1}\log_2(\frac{1}{0+1+1}))$$

1

Measuring conditional entropy on each of X2 attributes
For X2=1 branch

$$-(\frac{2}{2+1+0}\log_2(\frac{2}{2+1+0})+\frac{1}{2+1+0}\log_2(\frac{1}{2+1+0})+\frac{0}{2+1+0}\log_2(\frac{0}{2+1+0}))$$

0.918

For X2=0 branch

$$-(\frac{0}{0+1+2}\log_2(\frac{0}{0+1+2})+\frac{1}{0+1+2}\log_2(\frac{1}{0+1+2})+\frac{2}{0+1+2}\log_2(\frac{0}{0+1+2}))$$

0.918

$$H(Y|X1) = \frac{4}{6} * 1.5 + \frac{2}{6} * 1$$
$$H(Y|X1) = 1.333$$

$$H(Y|X2) = \frac{3}{6} * 0.918 + \frac{3}{6} * 0.918$$
$$H(Y|X2) = 0.918$$

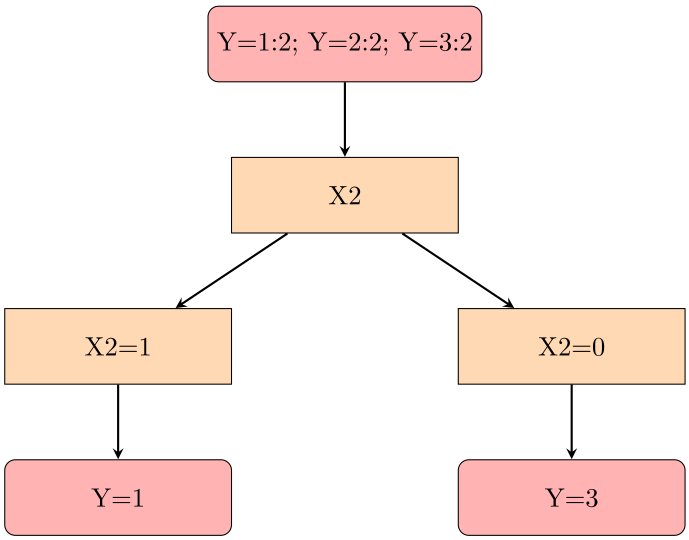
1.2

$$H(Y) = -(\frac{2}{2+2+2} * \log_2(\frac{2}{2+2+2}) + \frac{2}{2+2+2} * \log_2(\frac{2}{2+2+2}) + \frac{2}{2+2+2}) * \log_2(\frac{2}{2+2+2}))$$
$$H(Y) = -(\frac{1}{3} * \log_2(\frac{1}{3}) + \frac{1}{3} * \log_2(\frac{1}{3}) + \frac{1}{3} * \log_2(\frac{1}{3}))$$
$$H(Z) = 1.585$$

$$IG(X1) = H(Y) - H(Y|X1)$$
$$IG(X1) = 1.585 - 1.333$$
$$IG(X1) = 0.252$$

$$IG(X2) = H(Y) - H(Y|X2)$$
$$IG(X2) = 1.585 - 0.918$$
$$IG(X2) = 0.667$$

1.3



1.4

For X1=0 and X2=1 Since X2=1, we will take the first branch Then the tree will predict 1 (since it appears 2 times out of 3)