

1. Build the best decision tree using the following data. Use the Information Gain algorithm from class.

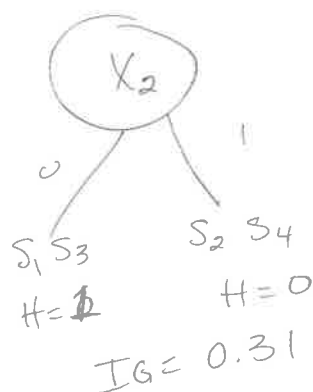
	$x_1$	$x_2$	label
$s_1$	0	0	0
$s_2$	0	1	1
$s_3$	1	0	1
$s_4$	1	1	1

Table 1: Training data.

$$H = -\frac{1}{4} \log \frac{1}{4} - \frac{3}{4} \log \frac{3}{4} = 0.5 + 0.31 = 0.81$$

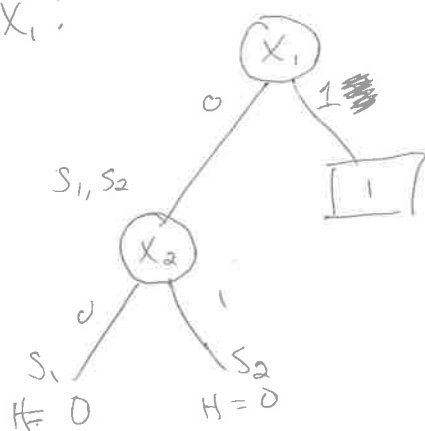


$$IG = 0.81 - \frac{1}{2} \cdot 1 - \frac{1}{2} \cdot 0 = 0.31$$

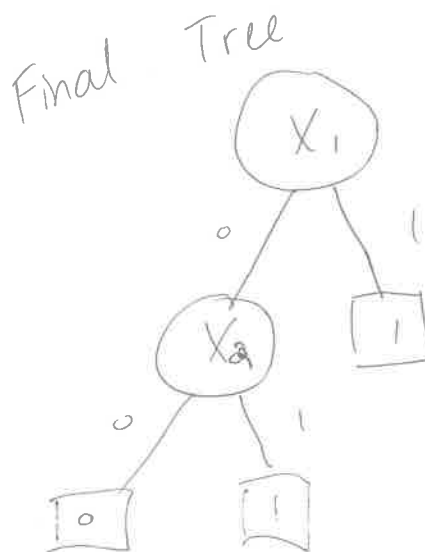


$$IG = 0.31$$

Choose  $x_1$ !



done



$p(c)$	$p(c) \log_2(p(c))$
$\frac{1}{4}$	-0.5
$\frac{1}{2}$	-0.5
$\frac{3}{4}$	-0.31
1	0

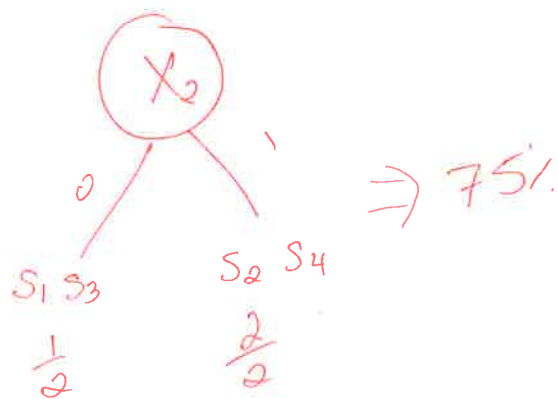
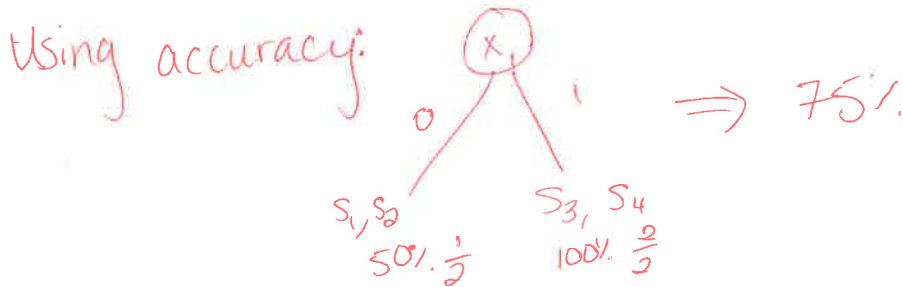
Table 2: Log table.

1. Show why Information Gain is a better measure of success than accuracy when building a decision tree for the following data.

$x_1$	$x_2$	label
0	0	0
0	1	1
1	0	1
1	1	1

Table 3: Training data.

Beginning Acc:  $75\%$   
 $\frac{3}{4}$



No tree would be built.  
 you would at least build  
 a depth-1 w/ infogain.  
 (see other side)

$$IG = H - \sum_{t \in T} p(t)H(t)$$

$$H = \sum_{c \in C} -p(c) \log_2(p(c))$$