Sec 2 BN 14

DataSet ⇒ P=8 N=6

$$E(S) = \frac{-8}{8+6}\log_2\left(\frac{8}{8+6}\right) - \frac{6}{6+8}\log_2\left(\frac{6}{6+8}\right)$$

$$= 0.985$$

For attributes

* early registeration

ER	P	N	Entropy
1	4	2	0.918
0	4	4	1

$$IG_1 = 0.985 - \frac{6}{14} * .918 - \frac{8}{14} * 1$$

Info gain|ER = 0.02

* finished homework

FH	P	N	Entropy
1	5	2	0.863
0	3	4	0.9892

$$\text{Info-gain}|_{FH} = 0.985 - \frac{7}{14} * 0.863 - \frac{7}{14} * 0.9892$$

$$= 0.061$$

* Senior

S	P	N	Entropy
1	5	3	0.954
0	3	3	1

$$\text{Info-gain}|_S = 0.985 - 0.954 * \frac{8}{14} - \frac{6}{14}$$

$$= 0.011$$

* likes coffee

LC	P	N	Entropy
1	3	1	0.811
0	5	5	1

$$\text{Info-gain}|_{LC} = 0.0041$$

Finished HW

$$E(S) = \frac{-5}{5+2}\log_2\left(\frac{5}{5+2}\right) - \frac{2}{2+5}\log_2\frac{2}{2+5}$$

$$= 0.8631$$

* left branch

5 2

attributes

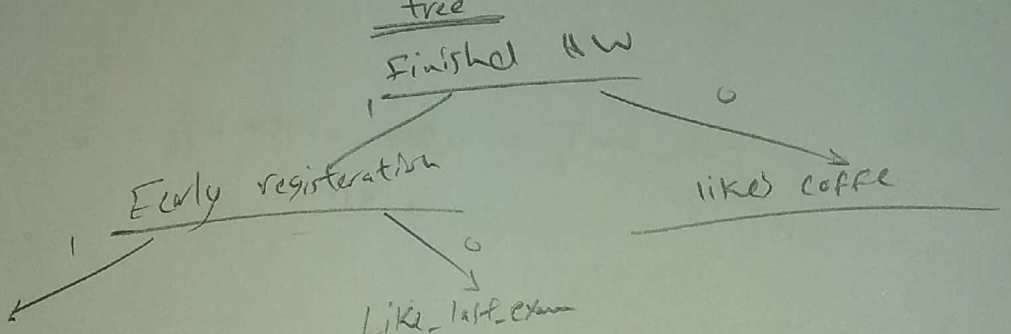
		+ve	-ve	entropy	
ER	1	3	0	0	Info-gain|ER = 0.291
	0	2	2	1	
S	1	3	2	0.971	IG|S = 0.169
	0	2	0	0	
LC	1	4	1	0.722	IG|LC = 0.0616
	0				
LLE	1	3	2	0.971	IG|LLE = 0.17
	0	2	0	0	

Finished HW right branch

$E(S) = \frac{-3}{3+4}\log_2\frac{3}{3+4} - \frac{4}{4+3}\log_2\frac{4}{4+3} = 0.9852$

		+ve	-ve	Entropy	
ER	1	1	2	0.918	$IG\|_{ER} = 0.02$
	0	2	2	1	
S	1	2	1	0.918	$IG\|_S = 0.811$
	0	1	3	0.811	
LC	1	2	0	0	$IG\|_{LC} = 0.469$
	0	1	4	0.772	
LLE	1	2	2	1	$IG\|_{LLE} = 0.02$
	0	1	2	0.918	

tree



for left left branch

ER = 1

∴ it is leaf node

ER	FH	S	LC	LLE	A
1	1	0	0	1	1
1	1	1	0	1	1
1	1	1	0	0	1

for left right branch

$E(S) = 0$

attributes		+ve	-ve	Entropy	
FH	1	2	2	1	$IG\|_{FH} = 0$
	0	0	0	0	
S	1	1	2	0.918	$IG\|_S = 0.3115$
	0	1	0	0	
LK	1	1	1	1	$IG\|_{LK} = 0$
	0	1	1	1	
LLE	1	1	2	0.918	$IG\|_{LLE} = 0.3115$
	0	0	0		

right left branch

Like coffe has A = 1 ∴ it is leaf node

right right branch

$E(S) = 0.721$

attributes

		true	-ve	Entropy	
ER	1	0	2	0	IG = 0.17
	0	1	2	0.918	
S	1	1	1	1	IG = 0.321
	0	0	3	0	
LC	1	0	0	0	IG = 0
	0	1	4	0.722	
LIE	1	1	2	0.918	IG = 0.17
	0	0	2	0	

tree