Date: _____ no: _____

P = 8 N = 6 Total = 14

$$\text{Entropy}(S) = \frac{-P}{P+n}\log\left(\frac{P}{P+n}\right) - \frac{n}{P+n}\log\left(\frac{P}{P+n}\right)$$

$$= \frac{-8}{14}\log\left(\frac{8}{14}\right) - \frac{6}{14}\log\left(\frac{6}{14}\right) = 0.985$$

* Entropy for each attribute:-

→ Early registration

	P	n	Entropy
1	4	2	0.918
0	4	4	1

* Calculate Average Information Entropy:-

$$I(\text{early registration}) = \frac{P_1+n_1}{P+n}\text{Entropy}(1) + \frac{P_2+n_2}{P+n}\text{Entropy}(0)$$

$$= \frac{4+2}{14} \times 0.918 + \frac{8}{14} \times 1 = 0.965$$

* Calculate Gain

Gain = Entropy (S) − I(Attribute)

= 0.989 − 0.965 = **0.024**

→ Finished homework II :-

	P	n	Entropy
1	5	2	0.863
0	3	4	0.985

$$I(\text{Finished homework II}) = \frac{5+2}{14} \times 0.863 + \frac{3+4}{14} \times 0.985 = 0.924$$

Gain = 0.989 − 0.924 = **0.065**

ROX

Scanned with CamScanner

Date: no:

→ Senior :-

	P	n	Entropy
1	5	3	0.954
0	3	3	1

$I(\text{Senior}) = \frac{5+3}{14} \times 0.954 + \frac{6}{14} \times 1 = 0.974$

$\text{Gain} = 0.989 - 0.974 = \boxed{0.015}$

→ Likes Coffee :-

	P	n	Entropy
1	3	1	0.8113
0	5	5	1

$I(\text{Likes Coffee}) = \frac{3+1}{14} \times 0.8113 + \frac{10}{14} \times 1 = 0.946$

$\text{Gain} = 0.989 - 0.946 = \boxed{0.043}$

→ Liked the last homework

	P	n	Entropy
1	5	4	0.991
0	3	2	0.971

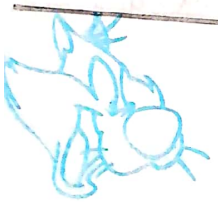
$I(\text{Liked the last homwork}) = \frac{5+4}{14} \times 0.991 + \frac{5}{14} \times 0.971$

$= 0.984$

$\text{Gain} = 0.989 - 0.984 = \boxed{0.005}$

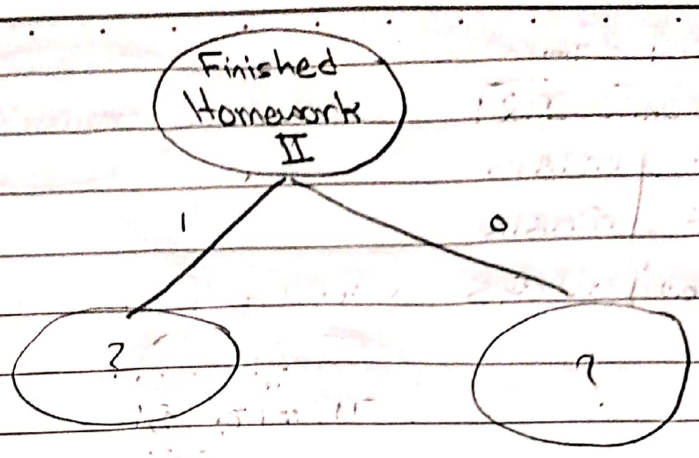
Attributes	Gain
Early registration	0.024
Finished homwork II	0.065
Senior	0.015
Likes coffee	0.043
Liked the last homework	0.005

∴ Finished homework II will be the root node as it has the highest gain.



Scanned with CamScanner

Date: _____ no: _____



for (1):

P = 5 N = 2 Total = 7

Entropy(S) = $-\frac{5}{7}\log\left(\frac{5}{7}\right) - \frac{2}{7}\log\left(\frac{2}{7}\right) = 0.863$

→ Early registration :-

	P	n	Entropy
1	3	0	0
0	2	2	1

I(Early registration) = $\frac{3}{7} \times 0 + \frac{2+2}{7} \times 1 = 0.571$

Gain = 0.863 - 0.571 = **0.292**

→ Senior :-

	P	n	Entropy
1	3	2	0.9709
0	2	0	0

I(Senior) = $\frac{5}{7} \times 0.9709 + 0 = 0.6935$

Gain = 0.863 - 0.6935 = **0.1695**

→ Likes Coffee :-

	P	n	Entropy
1	1	1	1
0	4	1	0.7219

I(Likes Coffee) = $\frac{2}{7} \times 1 + \frac{5}{7} \times 0.7219 = 0.8014$

Gain = 0.863 - 0.8014 = **0.0616**

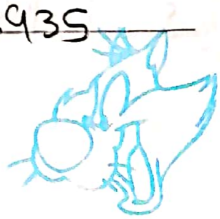
→ Likes the last homework

	P	n	Entropy
1	3	2	0.9709
0	2	0	0

I(LLH) = $\frac{5}{7} \times 0.9709 + 0 = 0.6935$

Gain = 0.863 - 0.6935

= **0.1695**

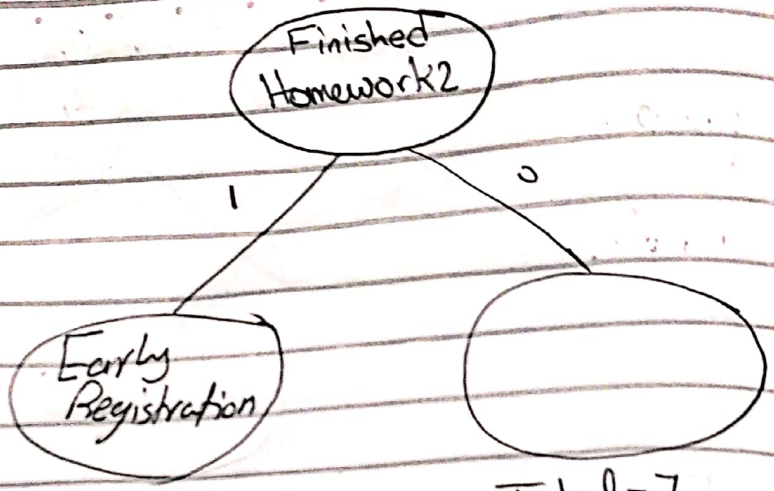


ROX

Scanned with CamScanner

Date: no:

Attribute	Gain
Early registration	0.292
Senior	0.1695
Likes Coffee	0.0616
Likes last homework	0.1695



for Finished Homework2 I = 0 ∴ P=3 n=4 Total=7

→ Early registration: Entropy(S) = 0.985

	P	n	Entropy
1	1	2	0.918
0	2	2	1

$I(ER) = \frac{3}{7} \times 0.918 + \frac{4}{7} \times 1 = 0.9648$

Gain = 0.985 − 0.9648 = **0.0202**

→ Senior

	P	n	Entropy
1	2	1	0.918
0	1	3	0.8113

$I(Senior) = \frac{3}{7} \times 0.918 + \frac{4}{7} \times 0.8113 = 0.857$

Gain = 0.985 − 0.857 = **0.128**

→ Likes Coffee

	P	n	Entropy
1	2	0	0
0	1	4	0.7219

$I(LC) = \frac{2}{7} \times 0 + \frac{5}{7} \times 0.7219 = 0.515$

Gain = 0.985 − 0.515 = **0.4**

→ Likes Last Homework

	P	n	Entropy
1	2	2	1
0	1	2	0.918

$I(LLH) = \frac{4}{7} \times 1 + \frac{3}{7} \times 0.918 = 0.9648$

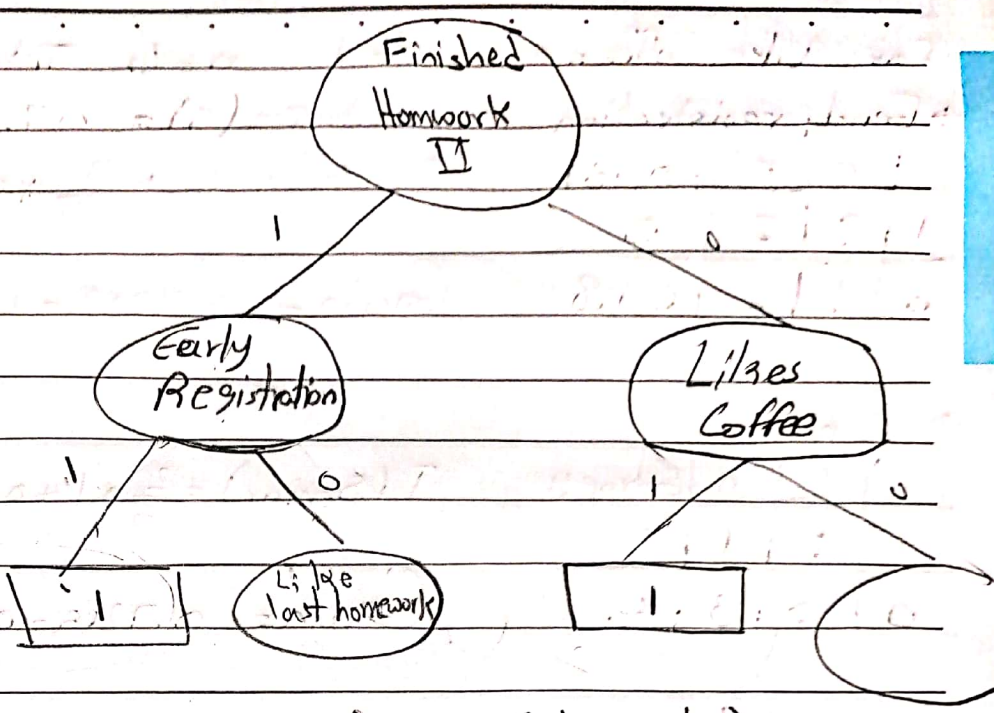
Gain = 0.985 − 0.9648 = **0.0202**

ROK

Scanned with CamScanner

Date: _____ no: _____

Attribute	Grain
ER	0.0202
Senior	0.128
LC	0.47
LLH	0.0202



for ER = 0 : P = 2 n = 2 Total = 4 Entropy(S) = 1

→ Senior

	P	n	Entropy
1	1	2	0.918
0	1	0	0

$I(\text{Senior}) = \frac{3}{4} \times 0.918 + 0 = 0.6885$

Grain = 1 − 0.6885 = 0.3115

→ Likes Coffee

	P	n	Entropy
1	1	1	1
0	1	1	1

$I(LC) = \frac{2}{4} \times 1 + \frac{2}{4} \times 1 = 1$

Grain = 1 − 1 = 0

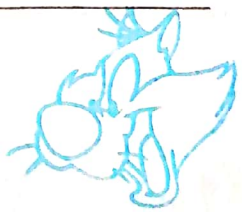
→ Likes last homework

	P	n	Entropy
1	1	2	0.918
0	1	0	0

$I(LLH) = 0.6885$

Grain = 1 − 0.6885 = 0.3115

Attribute	Grain
Senior	0.3115
LC	0
LLH	0.3115



Scanned with CamScanner

Date: _____ no: _____

For Like Coffee = 0 P = 1 n = 4 Total = 5

→ Early registration Entropy(S) = 0.7219

	P	n	Entropy
1	0	2	0
0	1	2	0.918

$I(ER) = 0 + \frac{3}{5} \times 0.918 = 0.5508$

Gain = 0.7219 − 0.5508 = **0.1711**

→ Senior

	P	n	Entropy
1	1	1	1
0	0	3	0

$I(Senior) = \frac{2}{5} \times 1 + 0 = 0.4$

Gain = 0.7219 − 0.4 = **0.3219**

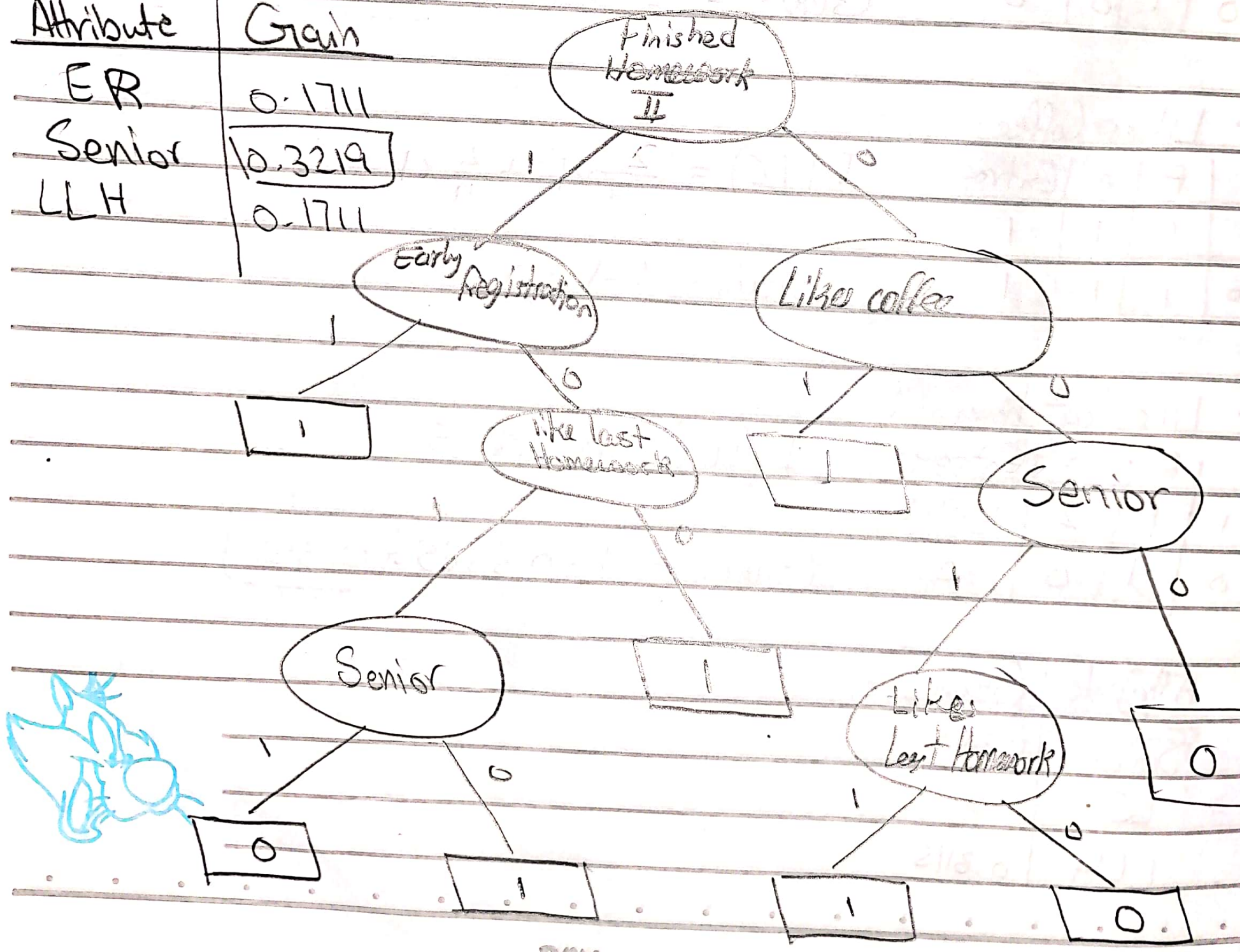
→ Likes Last homework

	P	n	
1	1	2	0.918
0	0	2	0

$I(ER) = \frac{3}{5} \times 0.918 = 0.5508$

Gain = **0.1711**

Attribute	Gain
ER	0.1711
Senior	0.3219
LLH	0.1711



Scanned with CamScanner

Date: _____ no: _____

For like last homework = 1 P = 1 n = 2 Total = 3 Entropy(S) = 0.918

Senior

	P	n	Entropy
1	0	2	0
0	1	0	0

I(Senior) = 0

Gain = 0.918 - 0 = 0.918

Likes coffee

	P	n	Entropy
1	1	1	1
0	0	1	0

$I(LC) = \frac{2}{3} = 0.67$

Gain = 0.918 - 0.67 = 0.248

for Senior = 1 P = 1 n = 1 Total = 2 Entropy(S) = 1

Early registration

	P	n	Entropy
1	0	0	0
0	1	1	1

$I(ER) = \frac{2}{2} \times 1 = 1$

Gain = 1 - 1 = 0

Like last Homework

	P	n	Entropy
1	1	0	0
0	0	1	0

I(LLH) = 0

Gain = 1 - 0 = 1

ROX

Scanned with CamScanner