

Decision Tree

A3

(1) Early Registration: values: 0, 1

$$S_{\text{total}} \Rightarrow [2, 6, 8] [8+, 6-]$$

$$S_{\text{total}} = -\frac{8}{14} \log_2 \frac{8}{14} - \frac{6}{14} \log_2 \frac{6}{14} = \underline{0.98}$$

$$S_0 \Rightarrow [4+, 4-]$$

$$S_0 = -\frac{4}{8} \log_2 \frac{4}{8} - \frac{4}{8} \log_2 \frac{4}{8} = \boxed{1}$$

$$S_1 \Rightarrow [4+, 2-]$$

$$S_1 = -\frac{4}{6} \log_2 \frac{4}{6} - \frac{2}{6} \log_2 \frac{2}{6} = \underline{0.918}$$

$$\text{Gain} = S_{\text{total}} - \sum_{\substack{\text{values} \\ \text{of feature}}} \frac{|S_v|}{|S|} \cdot S_v$$

$$\text{Gain}_{\text{early reg.}} = 0.98 - \frac{8}{14} \times 1 - \frac{6}{14} \times 0.918 = \boxed{0.015}$$

(2) Finished homework II: values: 0, 1

$$S_{\text{total}} = 0.98$$

$$S_0 \Rightarrow [3+, 4-]$$

$$S_0 = -\frac{3}{7} \log_2 \frac{3}{7} - \frac{4}{7} \log_2 \frac{4}{7} = \underline{0.985}$$

$$S_1 \Rightarrow [5+, 2-]$$

$$S_1 = -\frac{5}{7} \log_2 \frac{5}{7} - \frac{2}{7} \log_2 \frac{2}{7} = \underline{0.863}$$

$$\text{Gain}_{\text{finished H.W2}} = 0.98 - \frac{7}{14} \times 0.985 - \frac{7}{14} \times 0.863 = \underline{0.056}$$

(3) Senior: Values: 0, 1

$$S_{\text{total}} = 0.98$$

$$S_0 \Rightarrow [3+, 3-]$$

$$\underline{|S_0 = 1|}$$

$$S_1 \Rightarrow [5+, 3-]$$

$$S_1 = -\frac{5}{8} \log_2 \frac{5}{8} - \frac{3}{8} \log_2 \frac{3}{8} = \underline{|0.954|}$$

$$\cancel{\text{Gain}}_{\text{senior}} = 0.98 - \frac{6}{14} \times 1 - \frac{8}{14} \times 0.954 = \underline{|0.00628|}$$

(4) Likes Coffee: Values: 0, 1

$$S_{\text{total}} = 0.98$$

$$S_0 \Rightarrow [5+, 5-]$$

$$\underline{|S_0 = 1|}$$

$$S_1 \Rightarrow [3+, 1-]$$

$$S_1 = -\frac{3}{4} \log_2 \frac{3}{4} - \frac{1}{4} \log_2 \frac{1}{4} = \underline{|0.8113|}$$

$$\cancel{\text{Gain}}_{\text{Likes Coffee}} = 0.98 - \frac{10}{14} \times 1 - \frac{4}{14} \times 0.8113 = \underline{|0.034|}$$

(5) Liked last H.W: values: 0, 1

$$S_{\text{total}} = 0.98$$

$$S_0 \Rightarrow [3+, 2-]$$

$$S_0 = -\frac{3}{5} \log_2 \frac{3}{5} - \frac{2}{5} \log_2 \frac{2}{5} = \underline{|0.971|}$$

$$S_1 \Rightarrow [5+, 4-]$$

$$\underline{|S_1 = 0.99|}$$

$$\cancel{\text{Gain}}_{\text{liked last H.W.}} = 0.98 - \frac{5}{14} \times 0.971 - \frac{9}{14} \times 0.99 = \underline{|-0.0032|}$$

\therefore Root is "Finished last HW"

Last He Finished H.W. 2 = 0 branch

dataset:

Early regist.	Senior	Likes Coffee	Liked last HW	A
0	1	0	0	0
0	1	1	1	1
1	0	0	1	0
0	1	0	1	1
1	0	0	0	0
0	0	0	1	0
1	0	1	0	1

total entropy $S_{\text{total}} = [3+, 4-]$

$$\boxed{S_{\text{total}} = 0.985}$$

4) Liked last H.W.:

$$\frac{S_0}{\text{Early reg.}} = [1+, 2-]$$

1) Early reg.

$$S_0 \Rightarrow [2+, 2-]$$

$$\boxed{S_0 = 1}$$

$$S_1 \Rightarrow [1+, 2-]$$

$$\boxed{S_1 = 0.918}$$

$$S_0 \Rightarrow [1+, 2-]$$

$$\boxed{S_0 = 0.918}$$

$$S_1 \Rightarrow [2+, 2-]$$

$$\boxed{S_1 = 1}$$

$$\boxed{\text{Gain}_{\text{Liked last H.W.}} = 0.02}$$

$$\cancel{\text{Gain}_{\text{Early reg.}}} = 0.02$$

\therefore Next root is "Likes Coffee"

2) Senior:

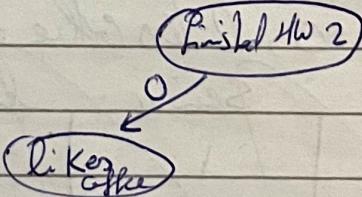
$$S_0 \Rightarrow [1+, 3-]$$

$$\boxed{S_0 = 0.811}$$

$$S_1 \Rightarrow [2+, 1-]$$

$$\boxed{S_1 = 0.918}$$

$$\cancel{\text{Gain}_{\text{Senior}}} = 0.128$$



3) Likes Coffee:

$$S_0 \Rightarrow [1+, 4-]$$

$$\boxed{S_0 = 0.722}$$

$$S_1 \Rightarrow [2+, 0-], S_1 = 0,$$

$$\cancel{\text{Gain}_{\text{Coffee}}} = 0.47$$

For the "Finished HW 2 → 0, likes coffee → 0 branch":

Early reg. Senior liked last HW. A

0	1	0	0
1	0	1	0
0	1	1	1
1	0	0	0
0	0	1	0

$$S_{\text{total}} \Rightarrow [1+, 4-] , S_{\text{total}} = 0.72$$

$$\Rightarrow \text{Early reg. : } S[1+, 2-] , S[0+, 2-]$$

$$S = 0.918$$

Finished H.W 2



likes coffee

Senior

A=1

$$\# \text{Gain}_{\text{early reg.}} = 0.169$$

$$2) \text{Senior's } S[0+, 3-] , S[1+, 1-]$$

$$S = 0$$

Senior

$$\# \text{Gain}_{\text{senior}} = 0.32$$

$$3) \text{Liked last HW, } S[0+, 2-] , S[1+, 2-]$$

$$S = 0$$

$$S = 0.918$$

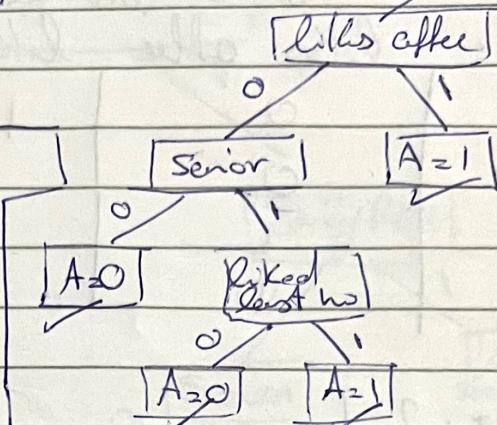
$$\# \text{Gain}_{\text{liked last HW}} = 0.169$$

"finished hw 2" likes coffee = 0, Senior = 1 "branches"

Early reg. liked last hw A

finished HW 2

"finished hw 2" = 1 "branch":



Early reg.	Senior	likes Coffee	liked the last hw	A
1, 1	0	0	1	1
1	1	0	1	1
0	1	0	1	0
0	1	0	0	1
0	0	1	1	1
1	1	0	0	1
0	1	1	1	0

total S $\Rightarrow [5+, 2-]$ $S_{\text{total}} = 0.863$

1) Early reg: $S_0[2+, 2-]$ $S_1[3+, 0-]$ $S_0[2+, 0-]$ $S_1[3+, 2-]$

$S_0 = 1$, $S_1 = 0$

4) liked last hw:

$S_0[2+, 0-]$ $S_1[3+, 2-]$

$S_0 = 0$, $S_1 = 0.97$

* Gain_{early reg} = 0.292

* Gain_{liked last hw} = 0.169

2) Senior: $S_0[2+, 0-]$ $S_1[3+, 2-]$

$S_0 = 0$, $S_1 = 0.97$

finished H.W. 2

* Gain_{senior} = 0.169

Early Reg. i
C
1

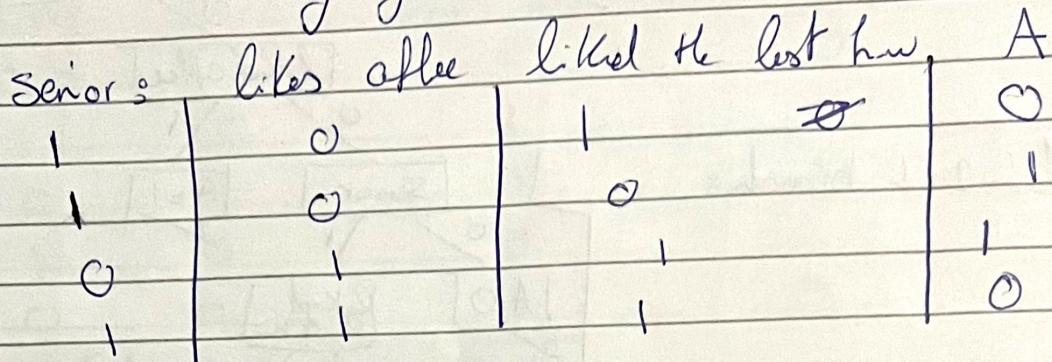
3) likes coffee: $S_0[4+, 1-]$ $S_1[1+, 1-]$

$S_0 = 0.722$ $S_1 = 1$

* Gain_{likes coffee} = 0.062

A=1

"finished H.W 2 = 1, Early reg. = 0" branch:



$$S_+ \Rightarrow [2+, 2-]$$

$$S_{\text{total}} = 1$$

1) Senior: $S_0 [1+, 0-], S_1 [1+, 2-]$

$$S_0 = 0, S_1 = 0.918$$

~~Gain~~
Senior = 0.3115

2) likes coffee: $S_0 [1+, 1-], S_1 [1+, 1-]$

$$S_0 = 1, S_1 = 1$$

~~Gain~~
likes coffee = 0

3) liked last hw: $S_0 [1+, 0-], S_1 [1+, 2-]$

$$S_0 = 0$$

$$S_1 = 0.918$$

~~Gain~~
liked last hw = 0.3115

Senior = 1 branch:

likes coffee liked last hw \Rightarrow A

0	1	0
0	0	1
1	1	0

Early Reg.

A = 1

Senior

A = 1

liked last hw

A = 1

A = 0

1) likes coffee $\Rightarrow S_0 [1+, 1-], S_1 [0+, 1-] \Rightarrow S_0 = 1, S_1 = 0$

2) liked last hw: $S_0 [1+, 0-], S_1 [0+, 2-]$

Final Tree :

