

Problem 2.

I:

Attribute 1	0	1
T	2	1
N	3	4

$$GINI(T_0) = 1 - \left(\frac{2}{3}\right)^2 - \left(\frac{1}{3}\right)^2 = \frac{4}{9}$$

$$GINI(N_0) = 1 - \left(\frac{3}{7}\right)^2 - \left(\frac{4}{7}\right)^2 = \frac{24}{49}$$

$$GINI(Attrib1) = \frac{4}{9} \times \frac{3}{10} + \frac{24}{49} \times \frac{7}{10} = 0.476 //$$

Attribute 2

	0	1
≤ small	2	2
> small	3	3

$$GINI : \frac{1}{2}$$

$$GINI(Attrib2) = \frac{1}{2} = 0.5 //$$

$$-- : \frac{1}{2}$$

	0	1
≤ medium	3	4
> medium	2	1

$$\frac{24}{49}$$

$$\therefore GINI = \frac{24}{49} \cdot \frac{7}{10} + \frac{4}{9} \times \frac{3}{10} = 0.476 //$$

$$\frac{4}{9}$$

Attrib 3

ordered: 5, 11, 23, 24, 30, 54, 82, 132, 175, 186

8 17 23.5 27 42 68 107 153.5 180.5

	0	1
≤ 8	1	0
> 8	4	5

$$\frac{40}{81}$$

$$GINI = \frac{4}{9} = 0.44 //$$

	0	1
≤ 17	1	1
> 17	4	4

$$\frac{1}{2}$$

$$GINI = 0.5$$

$$\frac{1}{2}$$

	0	1
≤ 23.5	1	2
> 23.5	4	3

$$\frac{4}{9}$$

$$GINI = 0.476.$$

$$\frac{24}{49}$$

	0	1	
≤ 27	1	3	$\frac{3}{8}$
> 27	4	2	$\frac{4}{9}$

$$GINI = 0.416$$

	0	1	
≤ 42	2	3	$\frac{12}{25}$
> 42	3	2	$\frac{12}{25}$

$$GINI = \frac{13}{25} = 0.48$$

≤ 68	2	4	$\frac{4}{9}$
> 68	3	1	$\frac{3}{8}$

$$GINI = 0.416$$

≤ 107	3	4	$\frac{24}{49}$
> 107	2	1	$\frac{4}{9}$

$$GINI = 0.476$$

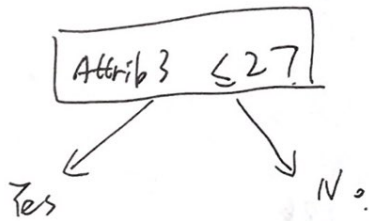
≤ 153.5	4	4	$\frac{1}{2}$
> 153.5	1	1	$\frac{1}{2}$

$$GINI = 0.5$$

≤ 180.5	4	5	$\frac{40}{81}$
> 180.5	1	0	0

$$GINI = 0.44$$

The smallest GINI Index is 0.416 from attrib 3 ≤ 27 .



II: Attrib 3 ≤ 27 :

→ Attrib 1.

0	0	0
1	3	$\frac{3}{8}$

GINI 0.375

→ Attrib 2.

	1	2	$\frac{4}{9}$
$\leq \text{small}$			
$> \text{small}$	0	1	0

GINI = 0.33

→ Attrib 3

pure Node at Attrib 3 ≤ 8

GINI = 0

→ smallest

II: Attrib 3 > 27 :

→ Attrib 1:

	2	1	$\frac{4}{9}$
Yes			
No	2	1	$\frac{4}{9}$

GINI = $\frac{4}{9} = 0.44$

→ Attrib 2:

	1	0	0
$\leq \text{small}$			
$> \text{small}$	3	2	$\frac{12}{25}$

GINI = $\frac{3}{5} = 0.4$

→ smallest

	0	1	
$\leq \text{medium}$	2	1	$\frac{4}{9}$
$> \text{medium}$	2	1	$\frac{4}{9}$

GINI = 0.44.

→ Attribute 3.

ordered	30	54	82	132	175	186
	42	68	107	153.5	180.5	

≤ 42	1	0	0
> 42	3	2	$\frac{12}{25}$

GINI = 0.4.

≤ 68	1	1	$\frac{1}{2}$
> 68	3	1	$\frac{3}{8}$

GINI = 0.416

≤ 107	2	1	$\frac{4}{9}$
> 107	2	1	$\frac{4}{9}$

GINI = 0.44.

≤ 153.5	3	1	$\frac{3}{8}$
> 153.5	1	1	$\frac{1}{2}$

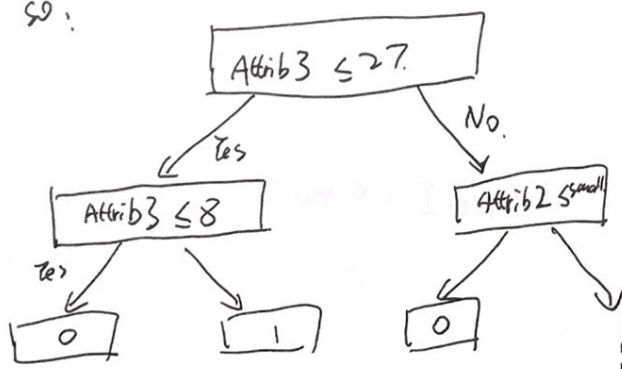
GINI = 0.416.

≤ 180.5	2	3	$\frac{12}{25}$
> 180.5	1	0	0

GINI = 0.4.

~~The lowest GINI is~~

so:



III: Attribute 3 > 27, Attrib 2 > small.

30 54 132 175 186.
 42 93 153.5 180.5

→ Attribute 1:

Yes	1	1	$\frac{1}{2}$
No	2	1	$\frac{4}{9}$

$$GINI = 0.467.$$

→ Attribute 2:

≤ medium.	0	2	0
> medium.	2	1	$\frac{4}{9}$

$$GINI = \frac{4}{15} = 0.267.$$

smallest

→ Attribute 3:

≤ 42	1	0	0
> 42	2	2	$\frac{1}{2}$

$$GINI = 0.4.$$

~~Attrib 3~~

≤ 93	1	1	$\frac{1}{2}$
> 93	2	1	$\frac{4}{9}$

$$GINI = 0.467$$

~~Attrib 3~~

~~≤ 93~~

~~> 93~~

≤ 153.5	2	1	$\frac{4}{9}$
> 153.5	1	1	$\frac{1}{2}$

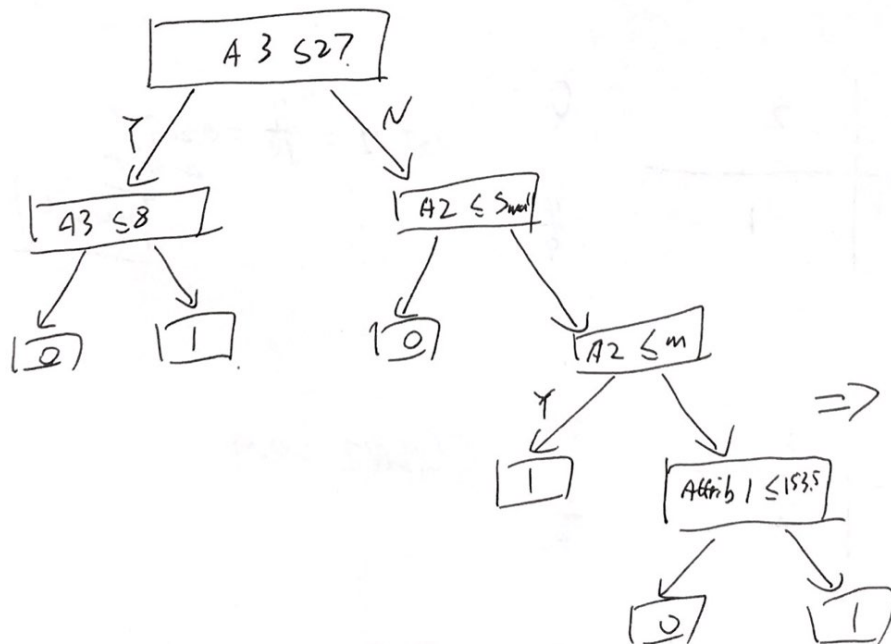
$$GINI = 0.467$$

≤ 180.5

> 180.5

2	2	0.5
1	0	0

$$GINI = 0.4$$



Attrib 3
 30, 132, 175
 class: 0 0 1