Classification Using CART algorithm

→ Our class variable is Caesarian (0=No, 1=Yes)

→ Independent Variables are –
- Age
- Heart Problem
- Blood Pressure
- Delivery Time
- Delivery Number

→ Total Items = 80

→ Gini for Class Variable

$$Gini(E) = 1 - \sum_{j=1}^{c} P_j^2$$

$$= 1 - \left[(34/80)^2 + (46/80)^2\right]$$

$$= 1 - [0.18 + 0.33]$$

$$= 0.49$$

Calculate Gini Gain for all variables.

Heart Problem	Caesarian 0	Caesarian 1		
0	28	22	apt	50
1	6	24	inept	30

Gini (E, Heart Problem) $= (50/80)$ gini(28,22)
$+ (30/80) \times$ gini(6,24)

$= (50/80)(1 - (28/50)^2 - (22/50)^2)$
$+ (30/80)(1 - (6/30)^2 - (24/30)^2)$

$= 0.308 + 0.12$

$= 0.428.$

Gini Gain (E, Heart Problem) $= 0.49 - 0.428 = \boxed{0.062}$

Gini Gain (E, Blood Pressure) $= 0.49 - 0.443 = 0.047$

Gini Gain (E, Delivery Time) $= 0.49 - 0.472 = 0.018$

Gini Gain (E, Delivery Number) $= 0.49 - 0.49 = 0.00$

Heart Problem

(Gini Gain = 0.062)

0 = apt

1 = inept

Gini Gain (apt, BP) = 0.49 - 0.48 = 0.01

Gini Gain (apt, DT) = 0.49 - 0.469 = 0.021

Gini Gain (apt, DN) = 0.49 - 0.461 = **0.029**

Gini Gain (inept, BP) = 0.49 - 0.241 = **0.249**

Gini Gain (inept, DT) = 0.49 - 0.315 = 0.175

Gini Gain (inept, DN) = 0.49 - 0.309 = 0.181