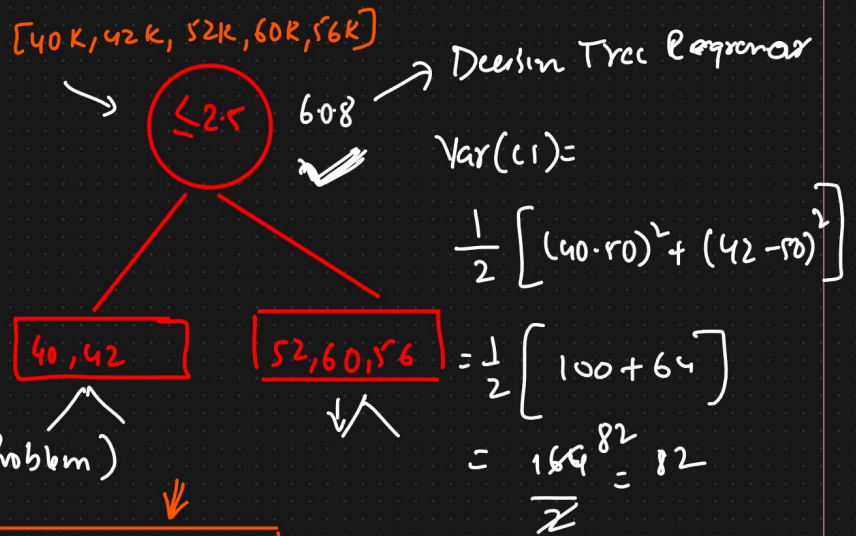
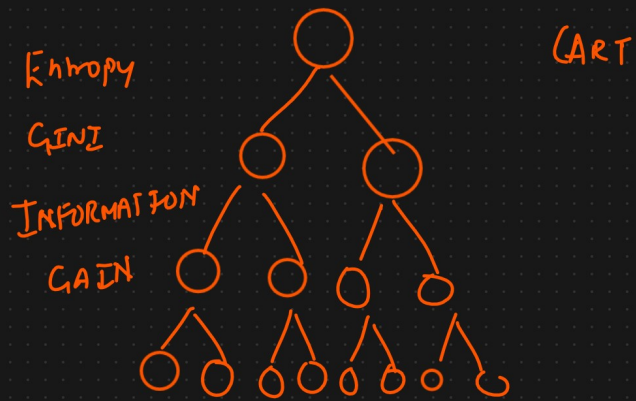
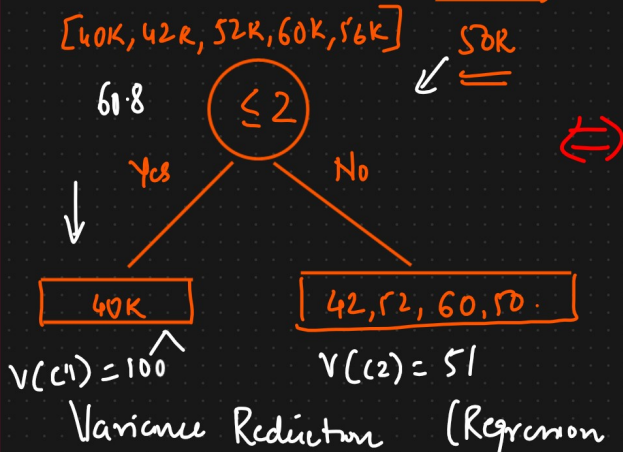


Decision Tree Regression

Exp	Career Gap	Salary
→ 2	Yes	40K
→ 2.5	Yes	42K
→ 3	No	52K
→ 4	No	60K
4.5	Yes	56K



$$\text{Variance} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$$

Mean Square Error

Average

Variance Reduction

V(c1) = 100

V(c2) = 51

$$\text{Variance (Root)} = \frac{1}{5} \left[(40-50)^2 + (42-50)^2 + (52-50)^2 + (60-50)^2 + (56-50)^2 \right]$$

$$= \frac{1}{5} [100 + 64 + 4 + 100 + 36]$$

$$= 60.8$$

Variance Reduction

V(c1) = 100

V(c2) = 51

Variance Reduction

V(c1) = 100

V(c2) = 51

$$\text{Variance (c1)} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$$

$$= \frac{1}{1} (40-50)^2$$

$$= 100$$

$$\text{Variance (c2)} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$$

$$= \frac{1}{4} \left[(42-50)^2 + (52-50)^2 + (60-50)^2 + (56-50)^2 \right]$$

$$= \frac{1}{4} [64 + 4 + 100 + 36]$$

$$= 51$$

Variance Reduction

$$\downarrow$$

$$= \text{Var}(\text{Root}) - \sum w_i \text{Var}(\text{child})$$

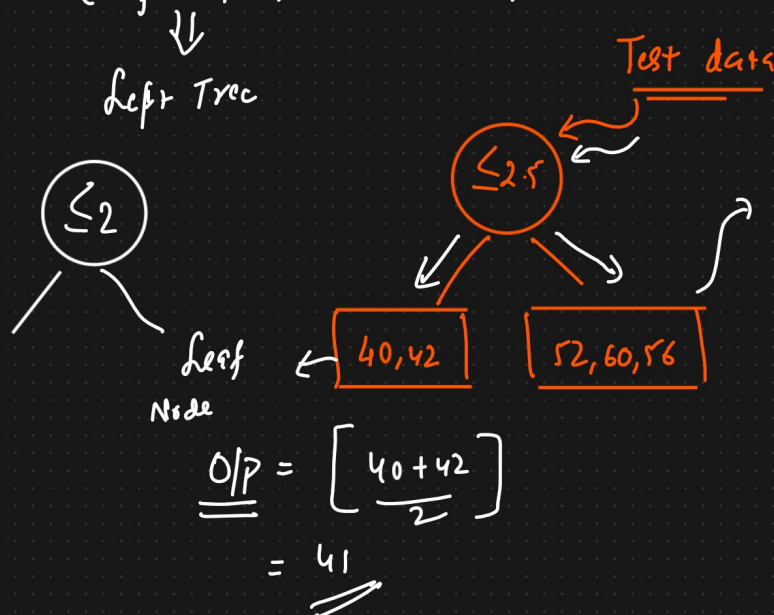
$$= 60.8 - \left[\frac{1}{5} \times 100 + \frac{4}{5} \times 51 \right]$$

$$= 60.8 - 20 - 40.8$$

Variance Reduction = 0

0 0.004

$$\text{Variance Reduction (Left Split)} < \text{VR (Right Split)}$$



Decision Tree
O/P Req.

$$\frac{52+60+56}{3} = 56$$