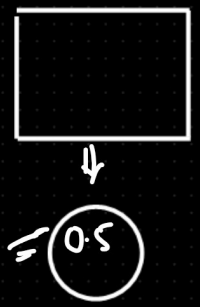
① XGBoost ⇒ Classification

Salary, Credit Score, Approval, Res1, Pred1, Res2

Salary	Credit Score	Approval	Res1	Pred1	Res2
<=50k	B	0	-0.5	-0.05	
<=50k	G	1	0.5		
<=50k	G	1	0.5		
>50k	B	0	-0.5		
>50k	G	-1	0.5		
>50k	N	i	0.5		
<=50k	N	0	-0.5		



Steps

1. Construct a base model.
2. Construct a Decision tree.
3. Similarity Weight
4. Calculate Gain

Residual ⇒ Actual − Predicted

(Approval) − 0.5 =

0.14 = [−0.5, 0.5, 0.5, −0.5, 0.5, 0.5, −0.5]

= Salary

<=5UK | >50K

[−0.5, 0.5, 0.5, −0.5] | [−0.5, 0.5, 0.5] 0.33

0

$$= \frac{(-0.5+0.5+0.5-0.5)^2}{0.5\times(1-0.5)+0.5\times(1-0.5)+0.5\times(1-0.5)+0.5\times(1-0.5)} = \frac{0}{1} = 0$$

(0.25 | 0.25 | 0.25 | 0.25)

LN = 0

RN = $\frac{(-0.5+0.5+0.5)^2}{0.5\times0.5+0.5\times0.5+0.5\times0.5} = \frac{(0.5)^2}{0.25+0.25+0.25}$

$= \frac{0.25}{0.75} = \frac{1}{3} = 0.33$

Similarity weight-

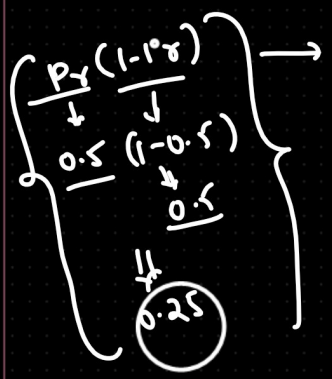
Current value = $\frac{\Sigma(\text{Residual})^2}{\Sigma Pa\times(1-Pa)+\lambda}$ $(\Sigma \text{Res})^2$

$Pa = 0.5$

Gain.

[left node + Right Node] − [Root Node]

Root Node = $\frac{(-0.5+0.5+0.5-0.5+0.5+0.5-0.5)^2}{0.25+0.25+0.25+0.25+0.25+0.25+0.25} = \frac{(0.5)^2}{1.75} = \frac{0.25}{1.75} = 0.14$

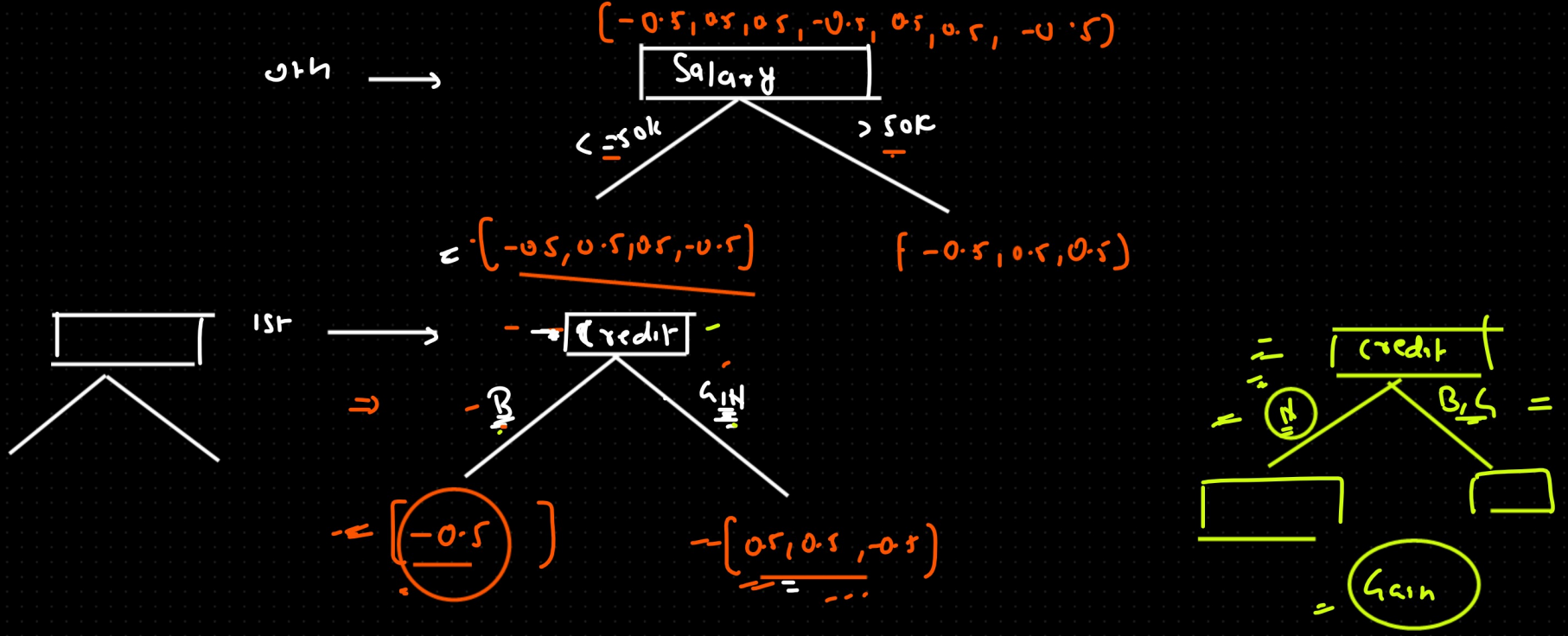


Left Node Similarity Score + RN S.C. − Root Node S.C.

↓

$(0 + 0.33) - 0.14 = 0.19$

=



$$(LN) = \frac{(-0.5)^2}{0.5(1-0.5)} = \frac{0.25}{0.5 \times 0.5} = \frac{0.25}{0.25} = 1$$

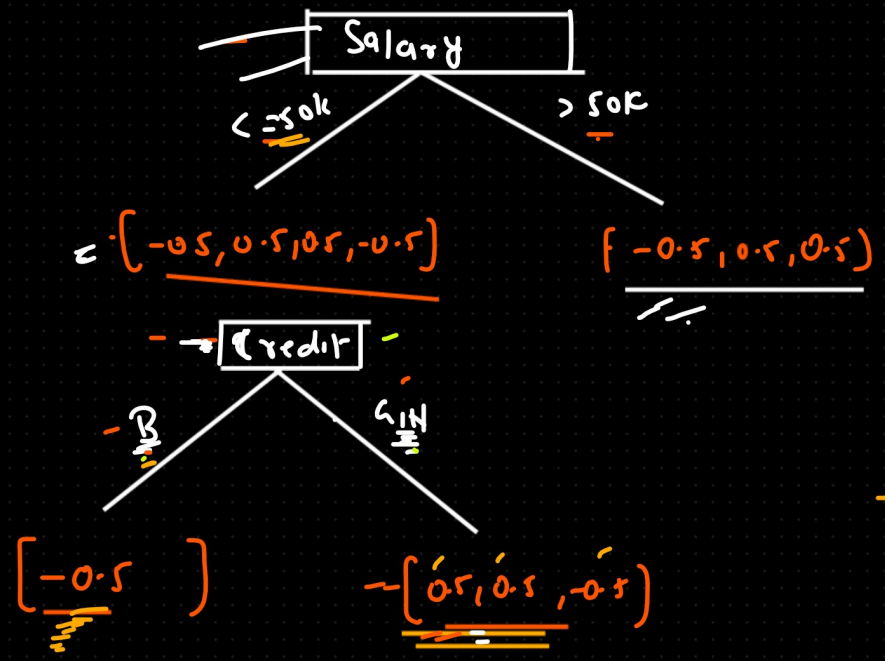
$$(RN) = \frac{(0.5 + 0.5 - 0.5)^2}{0.25 + 0.25 + 0.25} = \frac{0.25}{0.75} = \frac{1}{3} = 0.33$$

$$\text{Root node} = \frac{(-0.5 + 0.5 + 0.5 - 0.5)}{0.25 + 0.25 + 0.25 + 0.25} = \frac{0}{1} = 0$$

$$\text{Gain} = (LR + RN) - \text{Root Node}$$

$$= (1 + 0.33) - 0$$

$$= 1.33$$



$$Y\_Pred = M_0 + \alpha h(\theta)_1 + \alpha h(\theta)_2 + \cdots \alpha h(\theta)_n$$

$$Y\_Pred = 0.5 + DT_1 + \cdots\cdots + DT_N$$

test $\rightarrow$ $[\alpha = 50k,\ B] = ?$

$$\Rightarrow Y\_Pred = 0.5 + (-0.5) = 0$$

$$= \log\left(\frac{P}{1-P}\right) + \alpha\,(DT_1)$$

$\alpha = (0,1) \Rightarrow 0.1$

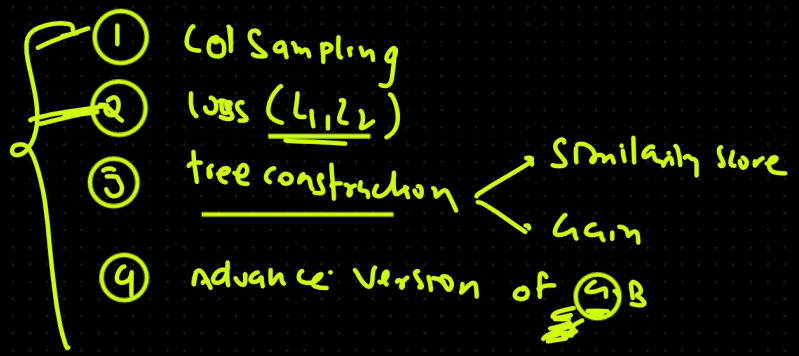
$$= \log\left(\frac{0.5}{1-0.5}\right) + (0.1)(-0.5)$$

$$= \log\left(\frac{0.5}{0.5}\right) + (-0.05)$$

$$= \log(1) + (-0.05)$$

$$= 0 - 0.05$$

$$= -0.05$$


1 Col Sampling
2 logs (L_1, L_2)
3 tree construction
Similarity score
Gain
4 advance version of 3