# Xgboost ML Algorithm Classification

## Dataset

| $x_1$ Salary | $x_2$ Credit | $y$ Approval | Error$(y-0.5)$ $\boxed{R_1}$ |
|---|---|---|---|
| <=50K | B | 0 | -0.5 |
| <=50K | G | 1 | 0.5 |
| <=50K | G | 1 | 0.5 |
| >50K | B | 0 | -0.5 |
| >50K | G | 1 | 0.5 |
| >50K | N | 1 | 0.5 |
| <=50K | N | 0 | -0.5 |

## Steps

1. Construct a base Model
2. Construct a Decision Tree with root node.
3. Calculate Similarity Weight

$$S.W = \frac{(\Sigma \text{Residuals})^2}{\Sigma Pr(1-Pr)}$$

4. Calculate Gain

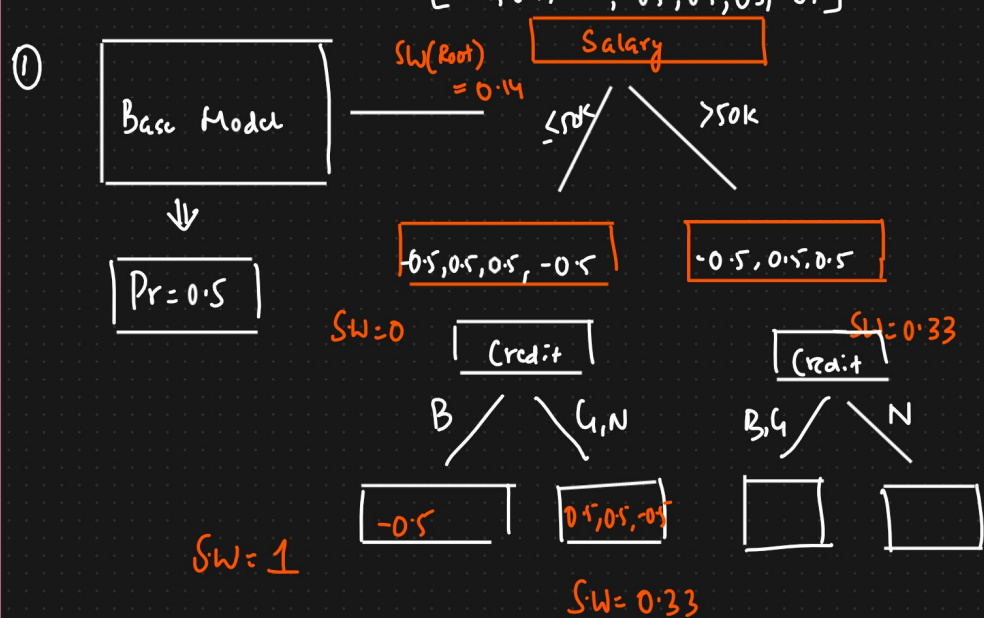
$$S.W(\text{Root}) = \frac{0.25}{1.75} = 0.14$$

$$\begin{array}{r} 0.25 \\ \times\ 7 \\ \hline 1.75 \end{array}$$

DT 1

$\{x_i, R_1\}$

$[-0.5, 0.5, 0.5, -0.5, 0.5, 0.5, -0.5]$

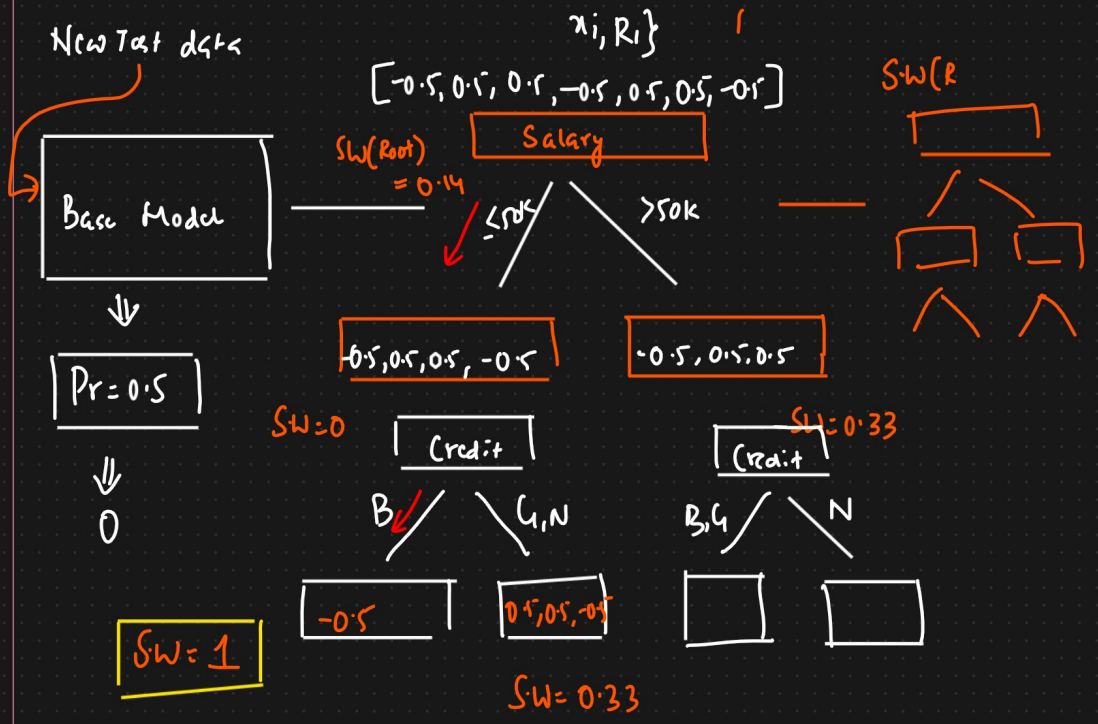



$$S.W(LC) = \frac{(\Sigma \text{residual})^2}{\Sigma Pr(1-Pr)} = \frac{[-\cancel{0.5}+\cancel{0.5}+\cancel{0.5}-\cancel{0.5}]}{0.5(0.5)+0.5(0.5)+0.5(0.5)+0.5(0.5)} = 0$$

$$SW(RC) = \frac{[-\cancel{0.5}+\cancel{0.5}+0.5]^2}{0.5(0.5)+0.5(0.5)+0.5(0.5)} = \frac{0.25}{0.75} = 0.33$$

$$\text{Gain} = 0 + 0.33 - 0.14 = 0.19$$

Final O/P  Classification




Test data

[≤50k, B] ⇒ O/P ??

log odds formula

$$\log(\text{odds}) = \log\left(\frac{P}{1-P}\right)$$

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

$$= 0 //$$

Test Data O/P $= \sigma\left(0 + \alpha(1)\right)$

$\alpha$ = learning Rate

$\alpha = 0.1$

$= \sigma(0 + 0.1)$

$= \sigma(0.1)$

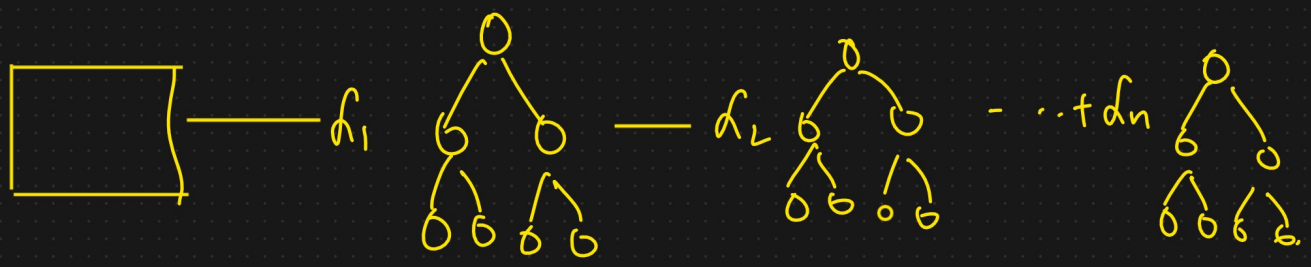
$\sigma = \frac{1}{1+e^{-z}}$

$= \frac{1}{1+e^{-0.1}}$

$= 0.52 //$ ⇒ Threshold ⇒ 0.6 ⇒ Domain Expert

$0.52 < 0.6 \Rightarrow 0$

# Xgboost Classifier




$$O/p = \sigma\left(\text{Base learner} + f_1(DT_1) + f_2(DT_2) + \cdots + f_n(DT_n)\right)$$

**Xg Boost classifier**