Xgboost Maths → Gradient Boosting

	(f1) Exp	(f2) Gap	(target) Salary	(Mean) P0	Res0	P1	Res1
①	↳ 2	1	40k	51k	-11k	46k	
②	* 2.5	1	42k	51k	-9k	46k	
③	3	0	52k	51k	1k	53.5k	
④	4	0	60k	51k	9k	—	
⑤	4.5	1	62k	51k	11k	—	

DT target → Res0

↓

Regression

Xgb Tree ≠ Decision Tree

Gain ↑↑

Similarity Score = $\frac{\sum (\text{Residuals})^2}{\text{num of Residuals} + \lambda}$

Hyperparameter

λ ←

Regularization Parameter

↓

Avoid Overfitting

* Gain = 60.5 + 28.8 − 0.16 = 89.14

$SS = \frac{(-11-9+1+9+11)^2}{5+1}$ ← Exp (-11, -9, 1, 9, 11)

$x <= 2$ | $x > 2$

$= 0.16$ (-11)

$SS = \frac{(-11)^2}{1+1}$

$= \frac{121}{2} = 60.5$

(-9, 1, 9, 11)

$SS = \frac{(-9+1+9+11)^2}{4+1}$

$= \frac{144}{5} = 28.8$

$SS = 0.16$

(-11, -9, 1, 9, 11)

Exp

$x = 2.5$ | 2.5

(-11, -9)

$SS = \frac{(-11-9)^2}{2+1}$

$= \frac{400}{3} = 133.33$

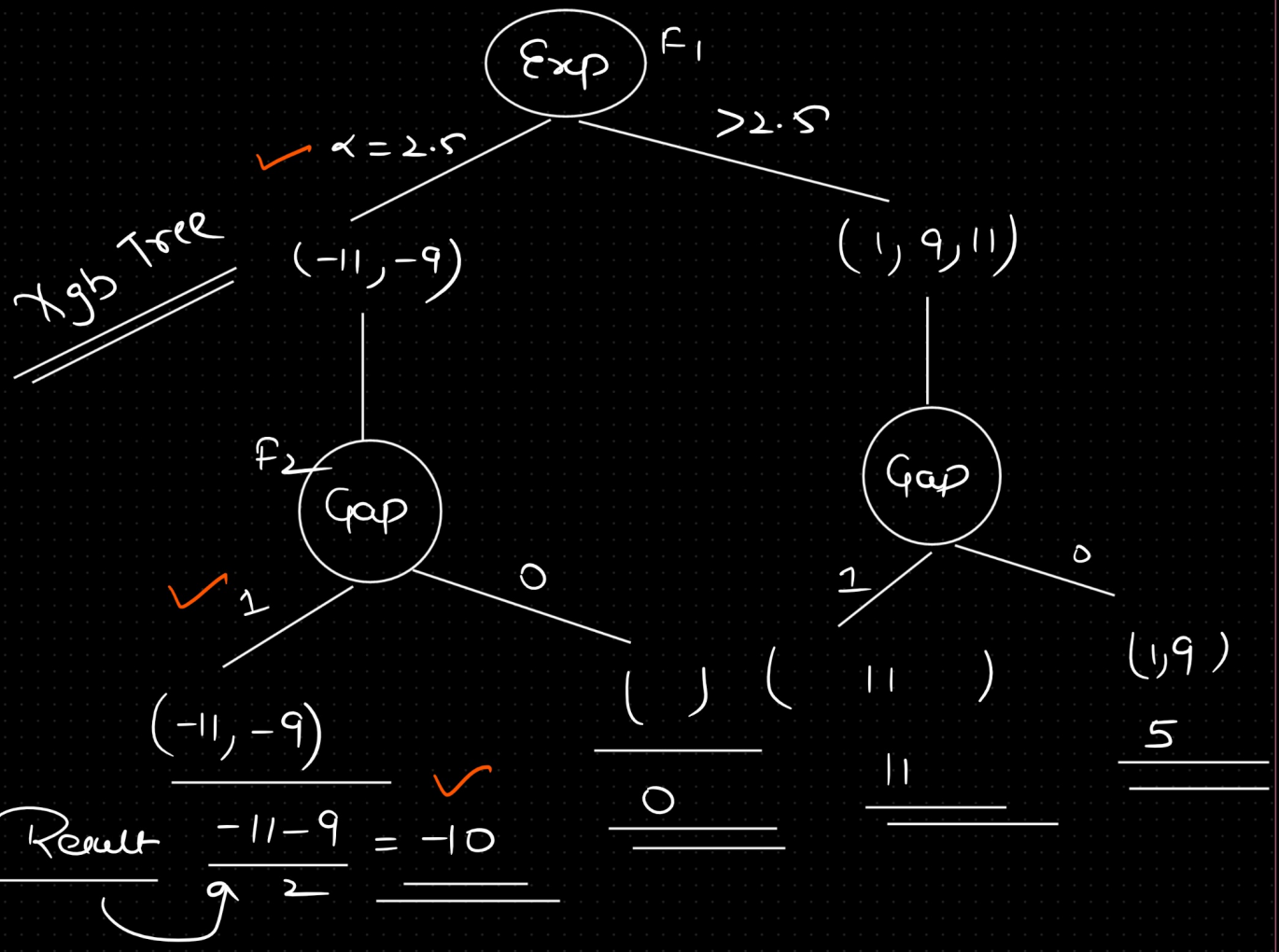
(1, 9, 11)

$SS = \frac{(1+9+11)^2}{3+1}$

$= 110.25$

Gain = 133.33 + 110.25 − 0.16

= 243.42



final Prediction $\longrightarrow$ $P_0 + \eta P_1 + \eta P_2 + \cdots$

$\eta = 0.5$ — Learning rate

$\Rightarrow 51 + (0.5) * (-10)$

$\Rightarrow 51 * \frac{5}{10} * (-10)$

$\Rightarrow 51 - 5$

$\Rightarrow 46\text{k}$

(3) $\Rightarrow$ $51 + (0.5) * 5$

$51 + \frac{5}{10} * 5$

$51 + 2.5$

$\Rightarrow 53.5$ K

Gamma $\rightarrow$ Hyperparameter

Gain - γ $\Rightarrow$ -ve $\rightarrow$ Pruning $\rightarrow$ to avoid the issue of Overfitting

Gain - γ $\Rightarrow$ +ve $\rightarrow$ no need to do Pruning

Gain - $\gamma < 0$

Gain $< \gamma \rightarrow$ Pruning

$\gamma \downarrow$ $\lambda \downarrow$ $\rightarrow$ Overfitted

Classification

$$SS = \frac{\left(\sum \text{residuals}\right)^2}{\sum \text{prob}(1-\text{prob}) + \lambda}$$

(f1) Age	(target) survived	P_0	Resi	P_1
28	0	0.5	-0.5	
34	0	0.5	-0.5	
17	1	0.5	0.5	
39	1	0.5	0.5	

$$\Rightarrow P_0 + \eta^{0.3} * (\text{o/p } xgb\ tree_1) + \_\_\_\_\_$$

$$\Rightarrow -0.6$$

Sigmoid function

$$\frac{e^{-0.6}}{1+e^{-0.6}} = 0.35$$

Probability

SS=0 $(0.5-0.5-0.5+0.5)/4*0.5(1-0.5)$

age < 22.5

y

n

$\lambda=0$

(0.5)

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

SS ⇒ 1

$(-0.5)\ -0.5, 0.5)$

$SS = \frac{[-0.5-0.5+0.5)^2}{3*0.5(1-0.5)}$

$=0.33$

Gain = 1 + 0.33 − 0

= 1.33