Gradient Boosting Algorithm

① Regression

② Classification

Regression Dataset

Exp	Degree	Salary (y)	$\hat{y}$	R_1 ($y-\hat{y}$)	R_2	$\hat{y}$	R_3	R_4
→ 2	B.E	50K	75K	-25K	-23	72.7	-22.7	-
3	Masters	70K	75K	-5K	-3	74.7	-4.7	-
5	Masters	80K	75K	5K	3	-	-	-
6	PHD	100K	75K	25K	20	-	-	-
		75K						

Steps

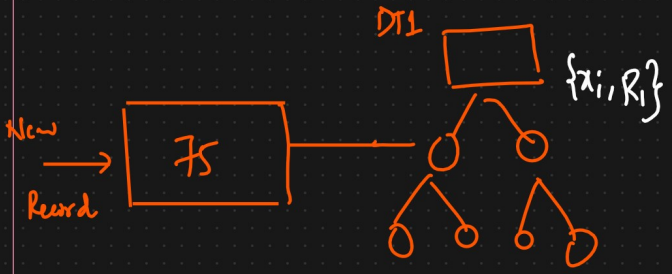
① Create a Base Model

Average = $\frac{50+70+80+100}{4} = 75$

② Compute Residuals, Error

③ Construct a Decision Tree

Consider inputs x_i and O/p R_i



Predicted O/p ⇒ 75 + (-23) = 75 - 23 = 52 {Overfitting}

Predicted O/P ⇒ $75 + \alpha(DT_1) = 75 + (0.1)(-23)$

α = learning (0 to 1)

$\alpha = 0.1$

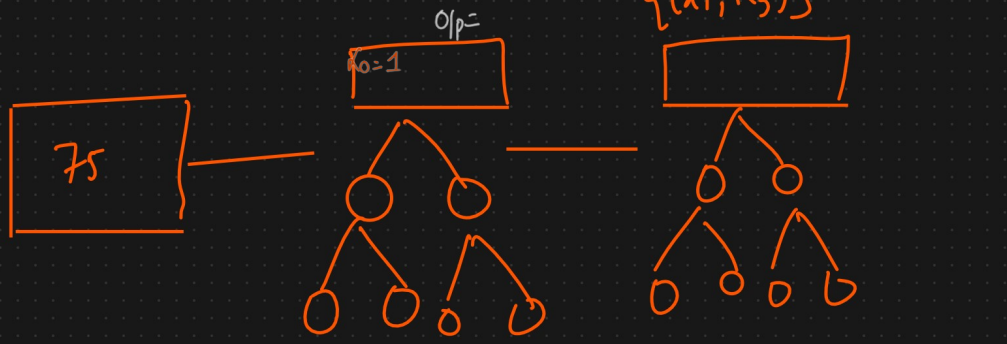
$= 75 - 2.3$

$= 72.7$

⇒ $75 + 0.1(-3)$

$= 75$ Base Model $= 74.7$

$= 75 + \alpha_1(DT_1) + \alpha_2(DT_2) + \cdots \alpha_n(DT_n)$



$$F(x) = \alpha_0 h_0(x) + \alpha_1(h_1(x)) + \alpha_2(h(x)) + \alpha_3(h(x)) \cdots + \alpha_n(h(x))$$

Learning Rate $\alpha = 0.1$

$$F(x) = \sum_{i=0}^{n} \alpha_i h_i(x)$$

⇒ Final Function of Gradient Boosting

Base Model

$= 75 + \alpha_1(DT_1) + \alpha_2(DT_2) + \cdots \alpha_n(DT_n)$

O/p=

$\alpha_0 = 1$