

GRADIENT BOOSTING ALGORITHM

① Regression

② CLASSIFICATION

Regression Dataset

Exp	Degree	Salary y	\hat{y}	$(y - \hat{y})$ R_1	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

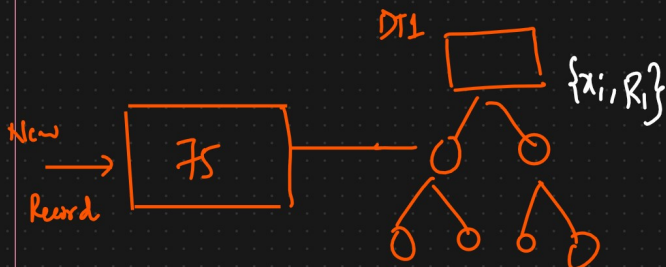
75

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

② Compute Residuals, Error

③ Construct a Decision Tree

Consider inputs x_i and o/p R_i



$$\text{Predicted o/p} \Rightarrow 75 + (-23) = 75 - 23 = \underline{\underline{52}} \quad \{\text{Overfitting}\}$$

Predicted o/p $\Rightarrow 75 + \alpha(DT_1) = 75 + (0.1)(-23)$

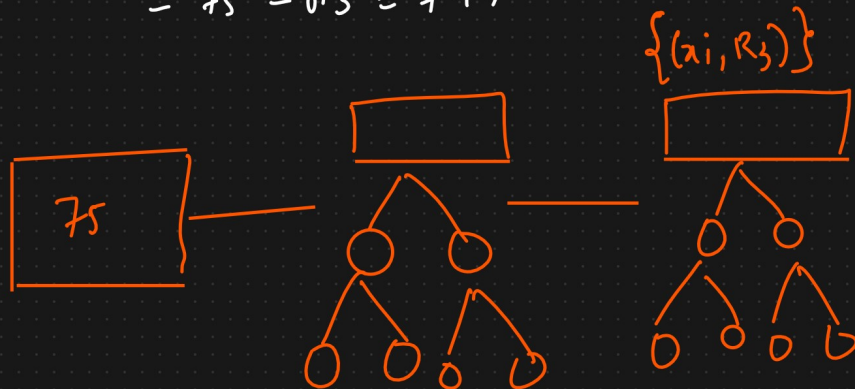
$\alpha = \text{learning rate}$
 $\alpha = 0.1$

$$= 75 - 2.3$$

$$= 72.7$$

$$\Rightarrow 75 + 0.1(-3)$$

$$= 75 - 0.3 = 74.7$$



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

Learning Rate $\alpha = 0.1$

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

\Rightarrow Final Function of Gradient Boosting