Gradient Boosting:

Dataset:

Exp (x_1)	Degree (x_2)	Salary (y)	$(y-\hat{y})$ R_1
2	B.E	50k	-25k
3	Masters	70k	-5k
5	~~PHD~~ Masters	80k	05k
6	PHD	100k	25k

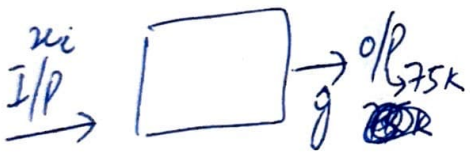
Avg. 75k ($\hat{y}$)

Suppose

Predicted R_2	$\hat{y}$	R_3
-23k	74.77	-24.77
-3k	74.97	-4.97
3k	-	-
20k	-	-

(A) Steps:

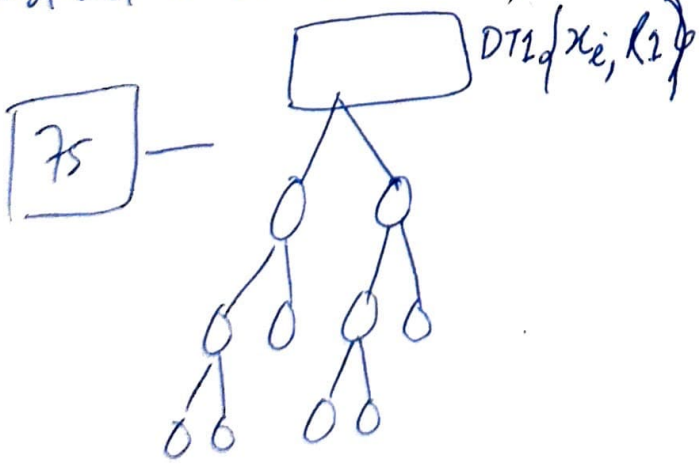
(1) Create a Base model



$$Avg = \frac{(50k + 70k + 80k + 100k)}{4} = 75k \Rightarrow \hat{y} = 75k$$

(2) Compute Residual ~~Error~~ and Error

(3) Construct a decision tree, Consider inputs x_i and o/p R_1



Predicted o/p = Base model + DT1

= 75 + (-23)

= (52) ⟶ {overfitting}

Predicted o/p for (x_1, y_1) = 75 + α(-23)

= 75 + (0.01)(-23)

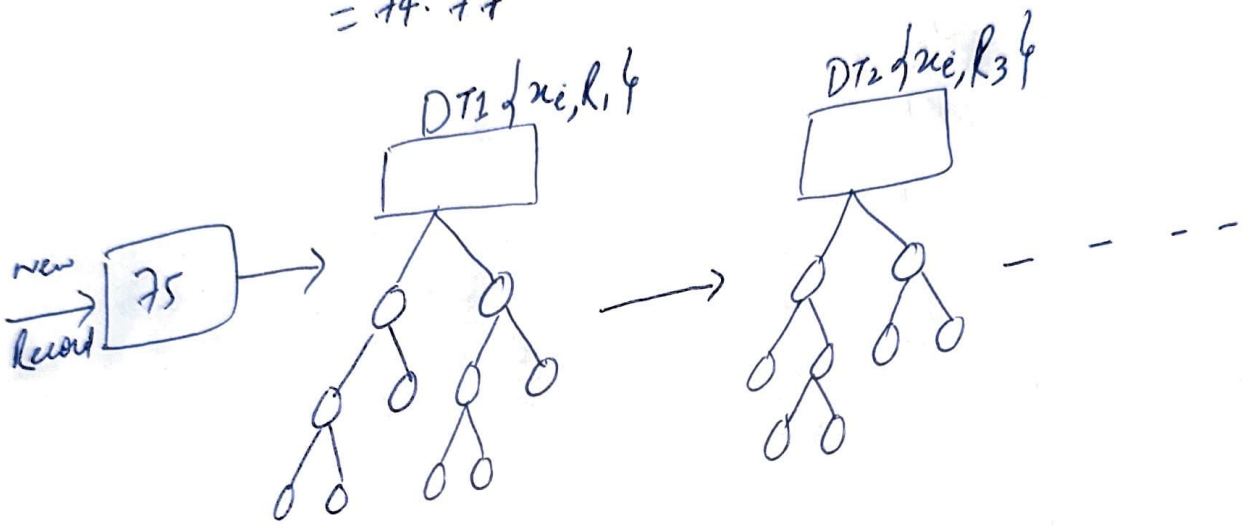
= 75 - 0.23

= 74.77

α = learning rate

~~$\alpha = 0.01$~~ $\alpha = \{0 \to 1\}$ large

$\alpha = 0.01$



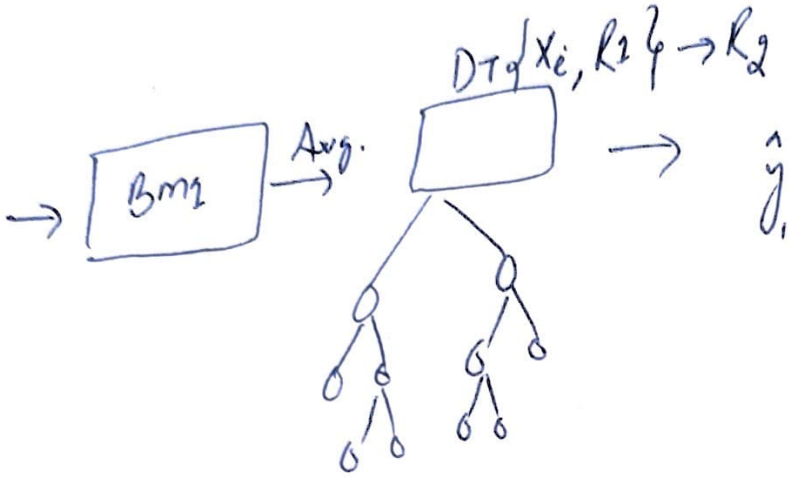
Mathematical formulation:

$$F(x) = \alpha_0 h_0(x) + \alpha_1 h_1(x) + \alpha_2 h_2(x) + \cdots \alpha_n h_n(x)$$

where, $\{\alpha_0, \alpha_1, \alpha_2 \cdots \alpha_n\}$ → learning rate $\{\alpha_0 = 1\}$ Ppt

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

X_1 Age	X_2 Bmi	③ Y Height	$\hat{y}_1$ Avg.	Residual(R_2)	R_2	$\hat{y}$	R_3
18	24	158	161.5	-3.5	-3	161.2	
20	26	160	161.5	-1.5	-1	161.4	
25	30	162	161.5	0.5	1	161.6	
30	40	166	161.5	4.5	4	161.9	



$$\hat{y}_1 = Bmi_1 + \alpha\{DT\{x_i, R_1\}\}$$
$$= 161.5 + 0.2\{-3\}$$
$$= 161.5 - 0.3$$
$$= 161.2$$

$$\hat{y}_2 = 161.5 + 0.1\{-1\}$$
$$= 161.5 - 0.1$$
$$= 161.4$$

$$\hat{y}_3 = 161.5 + 0.1\{1\}$$
$$= 161.6$$

$$\hat{y}_4 = 161.5 + 0.1\{4\}$$
$$= 161.9$$

$$F(x) = \alpha_0 h_0(x) + \alpha_1 h_1(x) + \alpha_2 h_2(x) + \cdots \alpha_n h_n(x)$$

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