# Gradient Boosting

→ It is a supervised machine learning algorithm which is used to solve both classification and regression problem by using boosting technique.

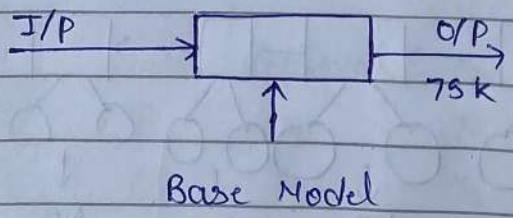
## * Working of Gradient Boosting :

Dataset :

| $x_1$ | $x_2$ | $y$ |
|---|---|---|
| Exp | Degree | Salary |
| 2 | B.E | 50k |
| 3 | Masters | 70k |
| 5 | Masters | 80k |
| 6 | PHD | 100k |

① Create a Base Model :

$$\text{Average of output} = \frac{[50k + 70k + 80k + 100k]}{4} = 75k$$




② Compute the Residual Errors :

$$\text{Residual Error} = y - \hat{y}$$

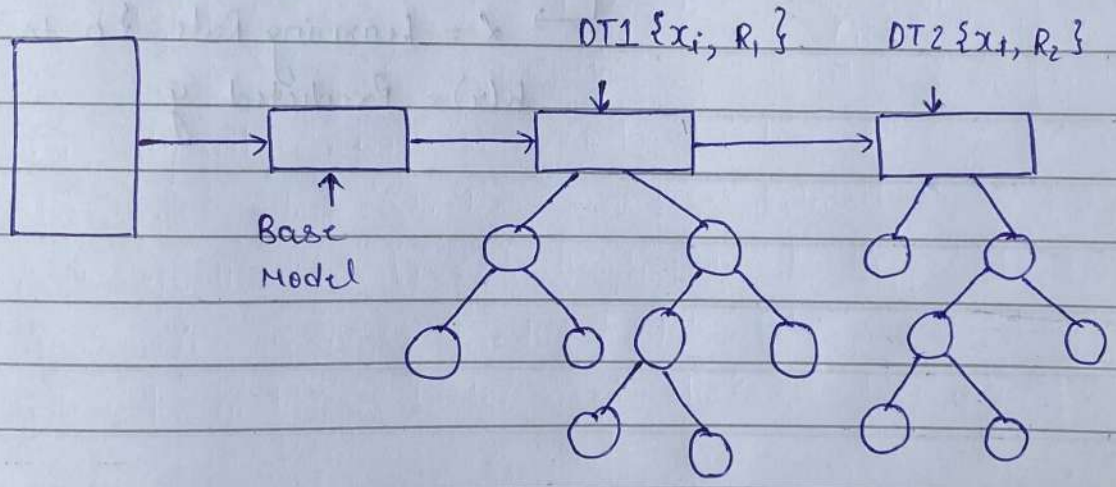
→ Average / Base Model Prediction

| Exp | Salary | $(y-\hat{y})$ $R_1$ | Predicted $R_2$ | $(y+R_2)$ $\hat{y}$ | $(y-\hat{y})$ $R_3$ |
|---|---|---|---|---|---|
| 2 | 50k | -25k | -23k | 74.77 | -24.77 |
| 3 | 70k | -5k | -3k | 74.97 | -4.97 |
| 5 | 80k | 5k | 3k | 75.03 | 4.97 |
| 6 | 100k | 25k | 20k | 75.2 | 24.8 |

$\bar{y} = 75k$

↳ Predicted from DT1

③ Construct a decision tree. Consider the input $x_1$ and output $R_2$ :




Predicted output : Base M o/p + DT o/p

= 75 + (-23) = 52 → very close {overfitting}

Therefore, Predicted output : BM o/p + $\alpha$ DT o/p

↳ learning Rate {0 to 1}

↓

0.01

| Records | Predicted output |
|---|---|
| 1 | 75 + 0.01 (-23k) = 75 - 0.23K = 74.77k |
| 2 | 75 + 0.01 (-3k) = 75 - 0.03k = 74.97k |
| 3 | 75 + 0.01 (3k) = 75 + 0.03k = 75.03k |
| 4 | 75 + 0.01 (20k) = 75 + 0.2k = 75.2k |

Note: This process of creating decision tree and predicting value then calculating residual error will keep on repeating until prediction become close or accurate.

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

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

$\alpha$ = learning Rate {0 to 1}

$h(x)$ = Predicted $y$