→ Boosting

→ Analogy

Anurag → Monu → Shobhit

3 problem solved

$e_0 = 10 - 3 = 7$

3 problem solved by Monu

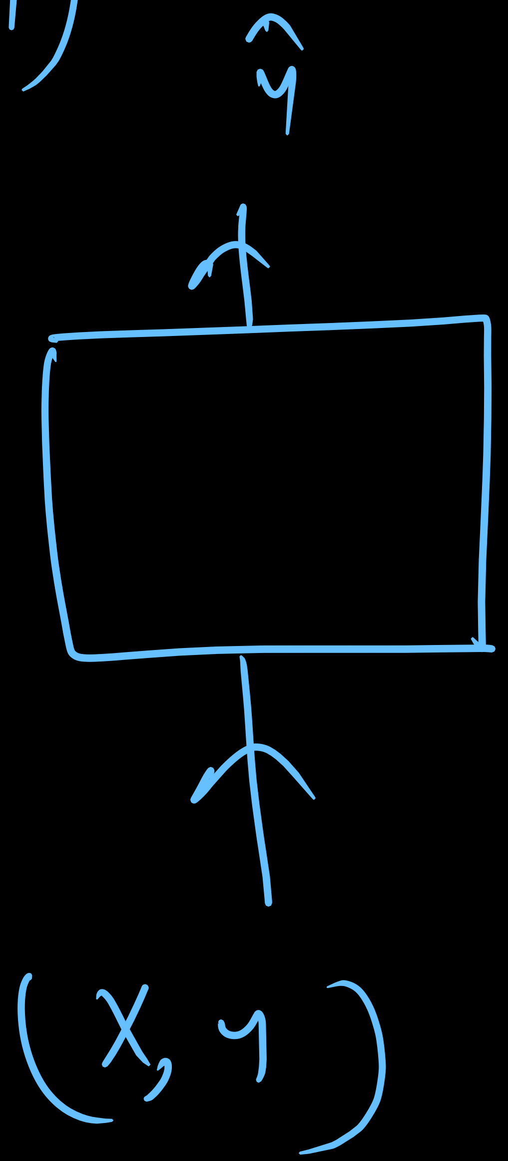
$e_1 = 10 - (3+3) = 4$

$e_2 = 10 - (3+3+4) = 0$

4 problem by Shobhit

→ Each friend is solving the unsolved question left by his/her previous friend

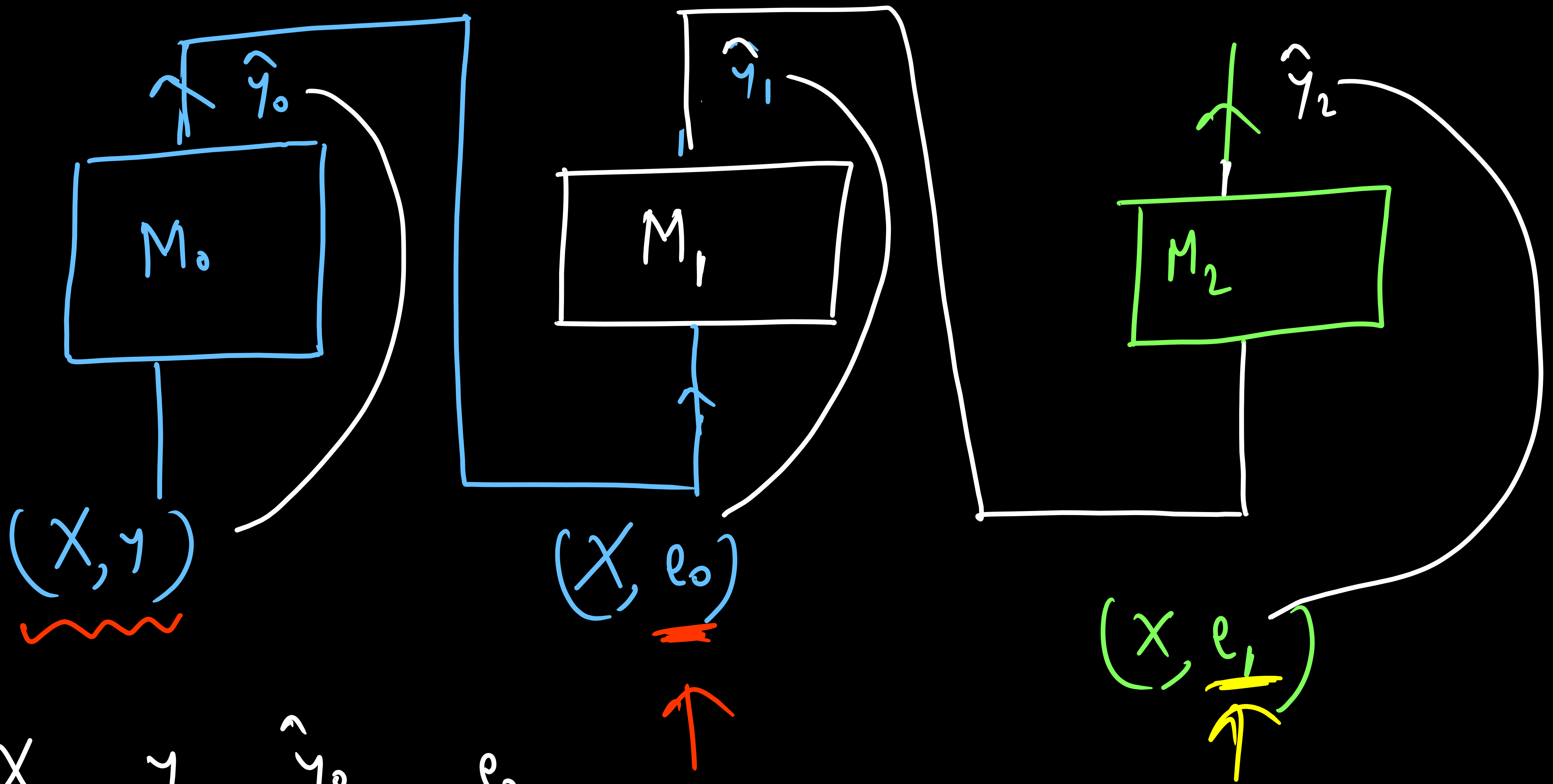
$e = (y - \hat{y})$



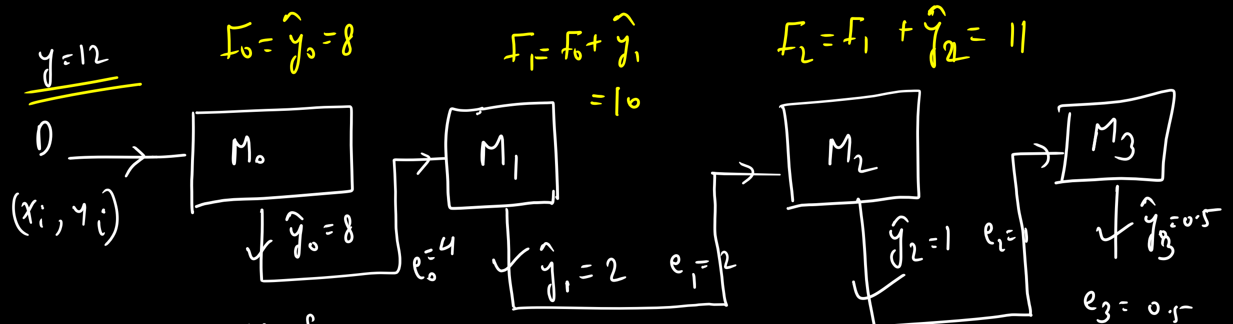
$e_0 = y - \hat{y}_0$

$e_1 = y - (\hat{y}_0 + \hat{y}_1)$

$e_2 = y - (\hat{y}_0 + \hat{y}_1 + \hat{y}_2)$



X	y	$\hat{y}_0$	e_0
	3.5	2	1.5
	3.8	2	1.8
	9	2.5	6.5

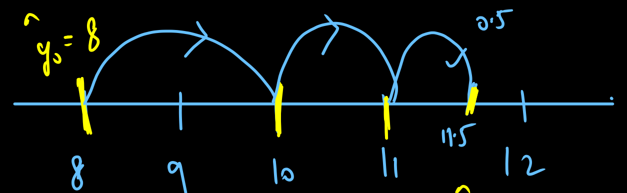


$e_0 = 12 - 8$
$= 4$

$e_1 = 12 - (8+2)$
$= 2$

$e_2 = 12 - (8+2+1)$
$= 1$

$F_2 =$



$F_2 = \hat{y}_0 + \hat{y}_1 + \hat{y}_2$
$= F_1 + \hat{y}_2$
$= 11$

$F_0 = \hat{y}_0$
$= 8$

$F_1 = \hat{y}_0 + \hat{y}_1$
$= F_0 + \hat{y}_1$
$= 10$

$y = 12$
$e_0 = 4$

$e_1 = 2$

$e_2 = 1$

F_2 low bias
& low variance

→ High bias + Low Variance

→ Low bias + low Variance

$F_0 = \hat{y}_0$

$F_1 = F_0 + \gamma \hat{y}_1$

$F_2 = F_1 + \gamma \hat{y}_2$
$= F_0 + \gamma (\hat{y}_1 + \hat{y}_2)$

$$\boxed{F_m = F_0 + \gamma \sum_{i=1}^{m} \hat{y}_i}$$

$\gamma = 0.1$
↳ Contribution of each of Base learner reduces

$F_0 + \gamma_1 \hat{y}_1 + \gamma_2 \hat{y}_2$

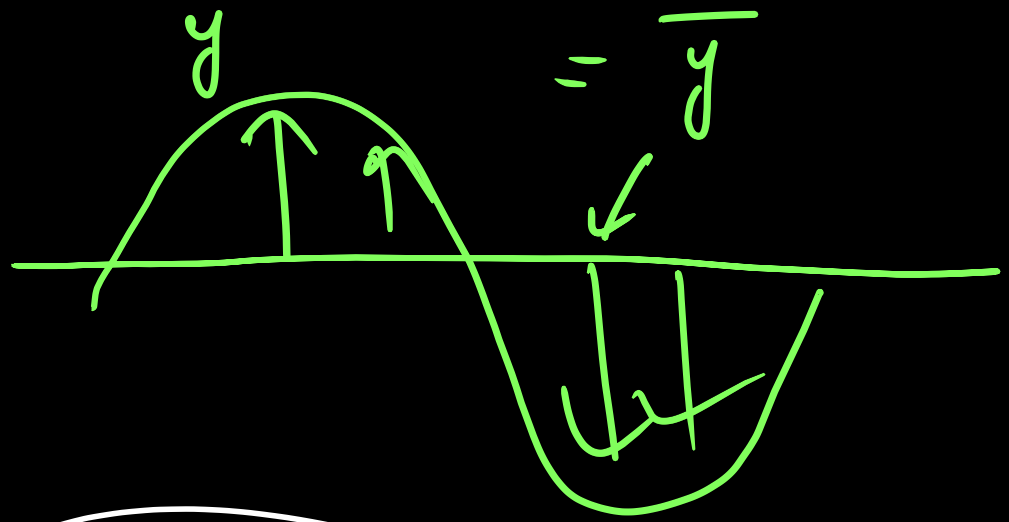
$F_0 + \gamma (\hat{y}_1 + \hat{y}_2)$

- Each model is called base learner or weak learner
- Initial base learner will have higher error
- Later base learner try to reduce the error made by previous base learner.

$M_0 \rightarrow$ Average Model

X	y
–	–
–	–
–	–
–	–

$\rightarrow \bar{y}$

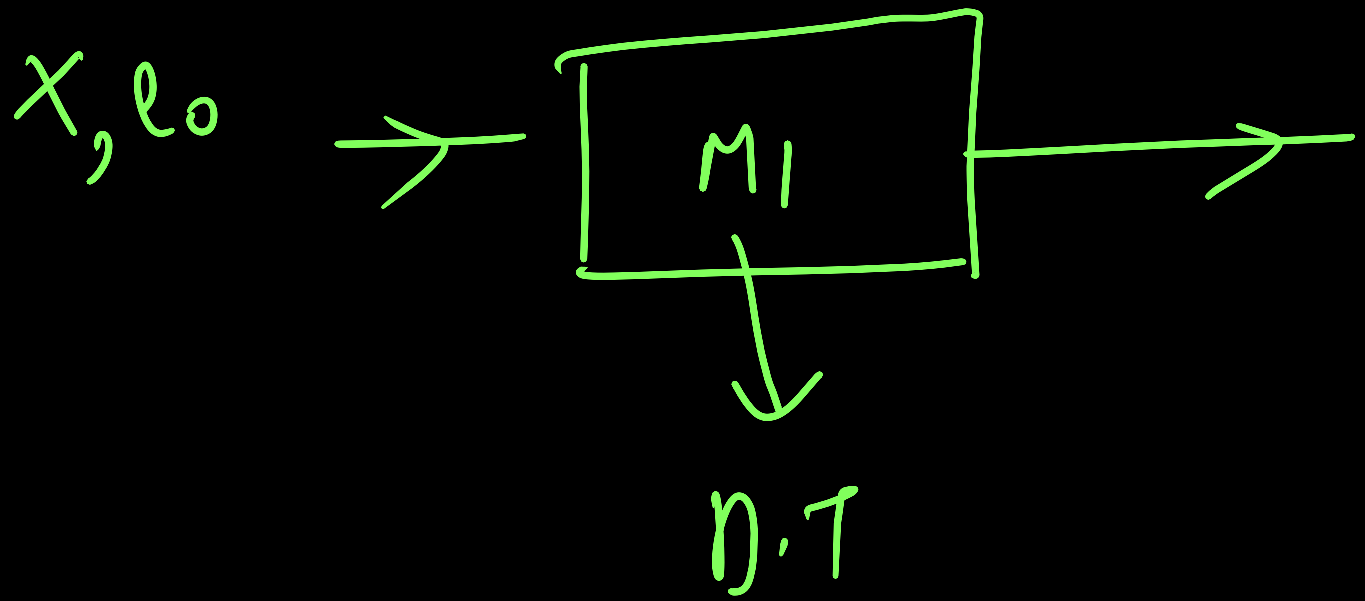


→ High bias & low variance

(Underfit)

$$e_0 = y - \hat{y}_0 = y - \bar{y}$$

M1



Underfit (High bias & Variance)

→ Break until 22:18

$$\overbrace{\qquad\qquad}^{X} \qquad \overbrace{\qquad}^{y}$$

Height	Gender	Weight(y)	$\hat{y}_0$ $= f_0(x)$	$e_0 = y - f_0(x)$
1.6	M	82	65.75	16.25
1.5	F	55	65.75	-10.75
1.4	F	61	65.75	-4.75
1.4	M	65	65.75	-0.75

$\longrightarrow M_0$

$$\hat{y}_0 = \frac{82 + 55 + 61 + 65}{4}$$

$$= 65.75$$

$$(X, e_0)$$

Height	Gender	Weight(y)	err_0
1.6	M	82	16.25
1.5	F	55	-10.75
1.4	F	61	-4.75
1.4	M	65	-0.75

F_0

65.75

e_0

Height	Gender	Weight(y)	err_0
1.6	M	82	16.25
1.5	F	55	-10.75
1.4	F	61	-4.75
1.4	M	65	-0.75

0.1

$F_1 = F_0 + \gamma \hat{y}_1$

$\hat{y}_1$	F_0	F_1	y	e_1
7.75	65.75	66.52	82	15.48
-7.75	65.75	64.97	55	-9.97
-7.75	65.75	64.97	61	-3.97
7.75	65.75	66.52	65	-1.52

F_0 - 0

M_1

$\boxed{M_0}$ +

65.75

Gender

M F

$$\frac{16.25 - 0.75}{2} = 7.75$$

$$\frac{-10.75 - 4.75}{2} = -7.75$$

e_0

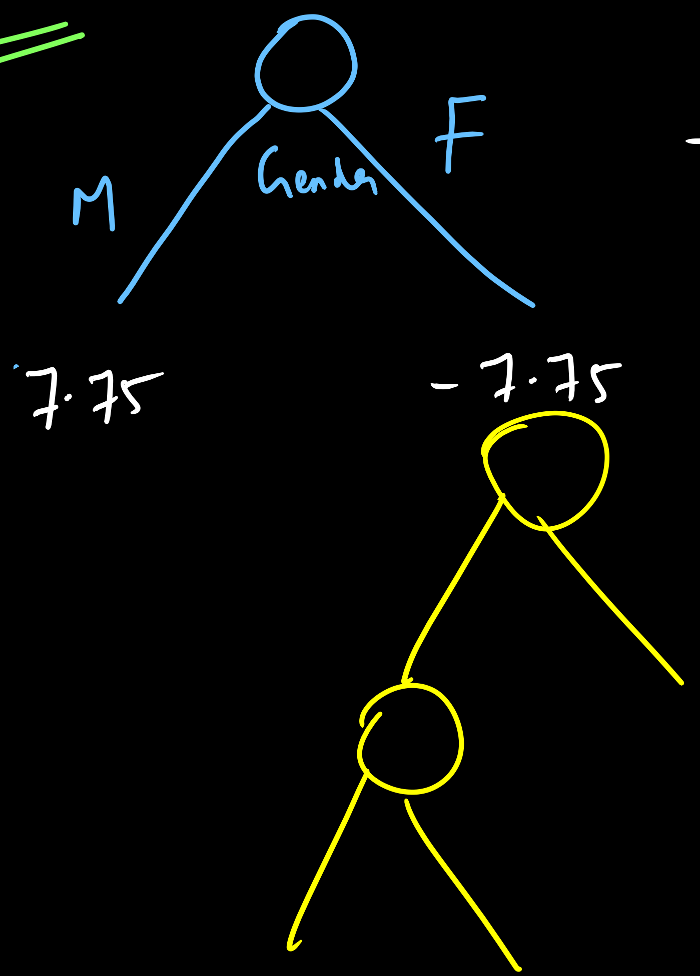
Height	Gender	Weight(y)	err_0
1.6	M	82	16.25
1.5	F	55	-10.75
1.4	F	61	-4.75
1.4	M	65	-0.75

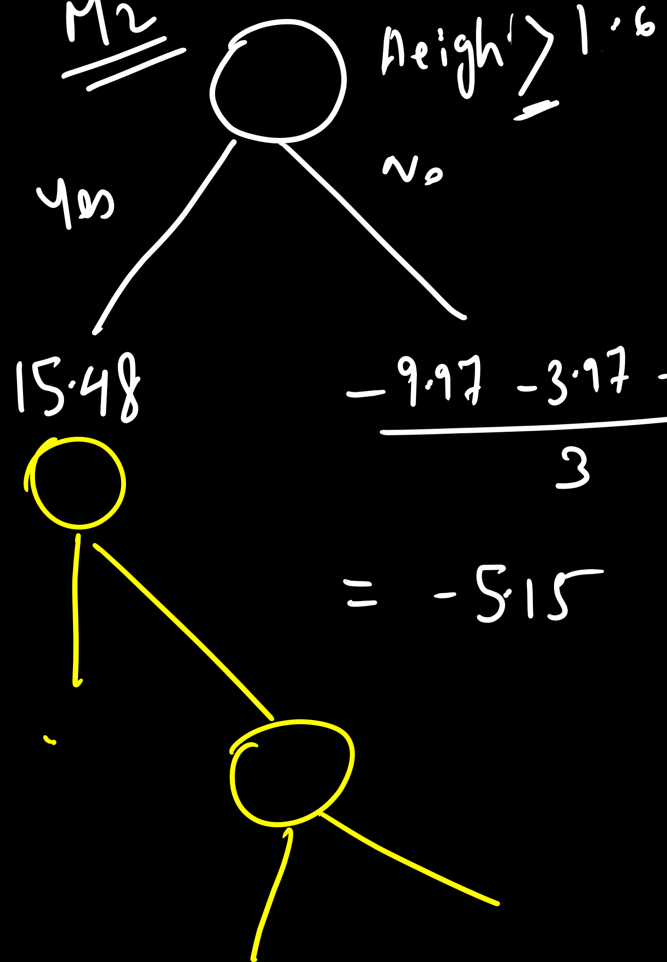
$F_1 = F_0 + \gamma \hat{y}_1$	y	e_1	$\hat{y}_2$	$F_2 = F_1 + \gamma \hat{y}_2$
66.52	82	15.48	15.48	
64.97	55	-9.97	-5.15	
64.97	61	-3.97	-5.15	
66.52	65	-1.52	-5.15	

$F_0 - 0$

$\boxed{M_0}$ + M_1 + M_2

65.75





$$\frac{-9.97 - 3.97 - 1.52}{3} = -5.15$$

DT

XGBoost

LightGBM

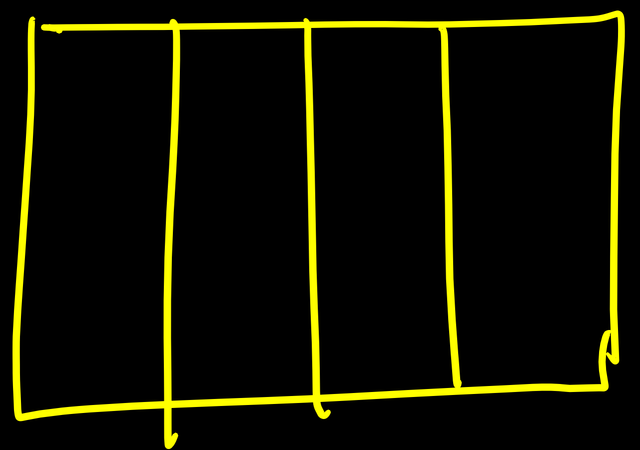
Catboost

RF

Gradient Boosting

Bagging

Tabular



Gradient Boosting Algorithm

1. Initialize model with a constant value:

$$F_0(x) = \underset{\gamma}{argmin} \sum_{i=1}^{n} L(y_i, \gamma)$$

2. for $m = 1\ to\ M$:

2-1. Compute residuals $r_{im} = -\left[\frac{\partial L(y_i,\ F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)}$ $\quad for\ i = 1,\ldots,n$

2-2. Train regression tree with features x against r and create terminal node reasions R_{jm} for $j = 1,\ldots,J_m$

2-3. Compute $\gamma_{jm} = \underset{\gamma}{argmin} \sum_{x_i \in R_{jm}} L(y_i, F_{m-1}(x_i) + \gamma) \quad for\ j = 1,\ldots,J_m$

2-4. Update the model:

$$F_m(x) = F_{m-1}(x) + \nu \sum_{j=1}^{J_m} \gamma_{jm} 1(x \in R_{jm})$$

$\boxed{M_0} \to \boxed{M_1} \to \boxed{M_2} \to \boxed{M_3}$

$y - \hat{y}_0 \to M_1$

$y - F_1 \to M_2$

T.A $\uparrow$

Test A $\uparrow$ $\rightarrow$ Best Fit $\leftarrow$ T.E $\downarrow$

Test E $\downarrow$

T.A $\uparrow$

Test A $\downarrow$ $\rightarrow$ Overfit $\leftarrow$ T.E $\downarrow$

Test E $\uparrow$

T.A $\downarrow$

Test A $\downarrow$ $\rightarrow$ Underfit $\leftarrow$ T.E $\uparrow$

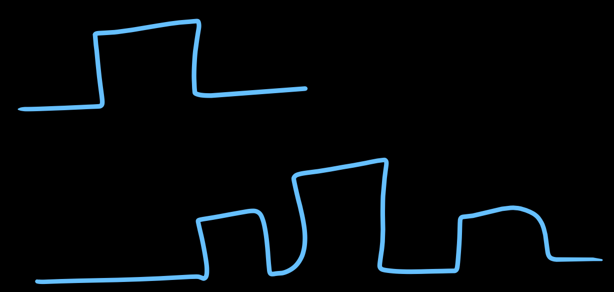
Test E $\uparrow$

Data drift

Test 1month $\longrightarrow$

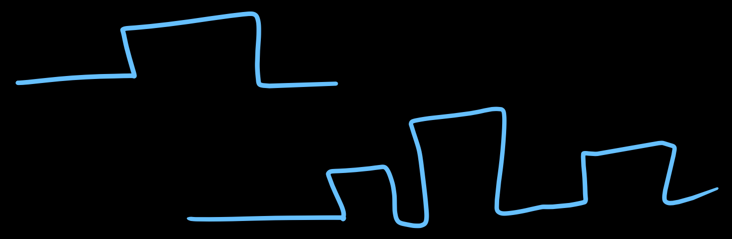
X	y
x_1, x_2	

Training



Test Current

Test



$$\boxed{y^{(i)} = \sum w_j x_j^{(i)} + w_0}$$

$w_j, w_0 \longrightarrow$ Normal

$$y^{(i)} = \sigma\left(\sum w_j x_j^{(i)} + w_0\right) \longrightarrow \text{Bernauli}$$

Height	Weight	y