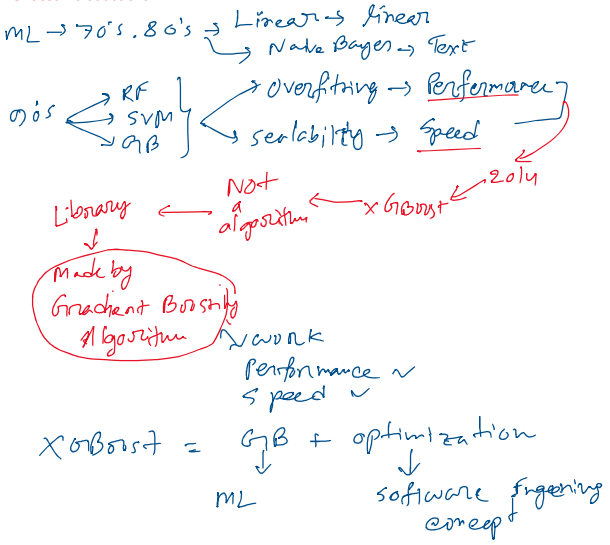
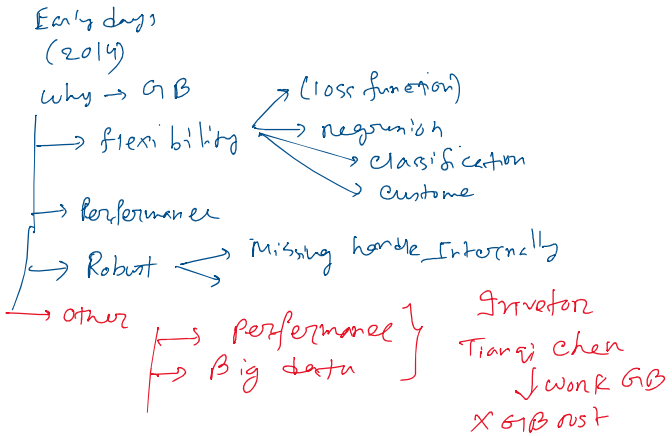
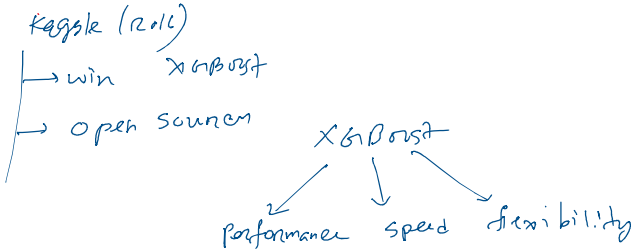
XGBoost Introductions



XGBoost History





Regression O/P

Ia | salary | model | res |

Avarage

Stage-1 Mean of O/P Features

Regression O/P

Actual − Pred

Ia	salary	model1	resi
6.7	4.5	7.3	−2.8
9.0	11.0	7.3	3.7
7.5	6.0	7.3	−1.3
5.0	8.0	7.3	0.7

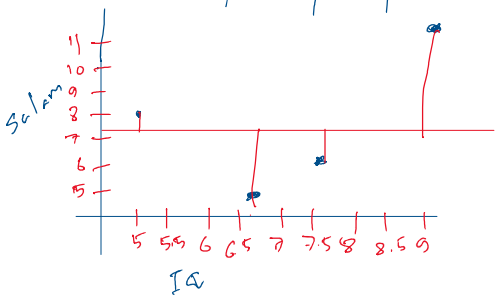
Stage−1 Mean of O/p Features

Stage−2 DT

I/p → Ia
O/p → resi

Spread DT for Ia

Ia	salary
6.7	4.5
9.0	11.0
7.5	6.0
5.0	8.0



Ia	resi
6.7	−2.8
9.0	3.7
7.5	−1.3
5.0	0.7

Stage−2 DT

$$\text{Similarity score} = \frac{(\text{Sum of residuals})^2}{\text{Number of residuals} + \lambda}$$

SS

Regularization parameter

−2.8, 3.7, −1.3, 0.7

$$SS = \frac{(-2.8+3.7-1.3+0.7)^2}{4+0} \qquad \lambda = 0$$

$$= \frac{(0.3)^2}{4} = \frac{0.09}{4} = 0.02 \quad SS_{root}$$

−2.8, 3.7, −1.3, 0.7

Ia
5.0] 5.85
6.7] 7.1
7.5] 8.25
9.0]

Ia	resi
5.0	0.7
6.7	−2.8
7.5	−1.3
9.0	3.7

① splitting 5.85 ✓ | ② Sp 7.1 ✗ | ③ Sp 8.25 ✓

① splitting criteria 5.85

Iq < 5.85

0.7 | -2.8, -1.3, 3.7

$SS_L = \frac{(0.7)^2}{1+0} = 0.49$

$SS_R = \frac{(-2.8-1.3+3.7)^2}{3+0}$

$= \frac{(-0.4)^2}{3}$

$= \frac{0.16}{3} = 0.05$

Gain = $(SS_L + SS_R) - SS_{Root}$

= 0.49 + 0.05 − 0.02

= 0.52

② Sp 7.1

Iq < 7.1

0.7, -2.8 | -1.3, 3.7

$SS_L = \frac{(0.7-2.8)^2}{2+0}$

$= \frac{4.41}{2}$

$= 2.20$

$SS_R = \frac{(-1.3+3.7)^2}{2}$

$= 2.88$

gain = $SS_L + SS_R - SS_{Root}$

= 2.20 + 2.88 − 0.02

= 5.06

③ Sp 8.25

Iq < 8.25

0.7, -2.8, -1.3 | 3.7

$SS_L = \frac{(0.7-2.8-1.3)^2}{3+0}$

$= 3.85$

$SS_R = \frac{(3.7)^2}{1+0} = 13.69$

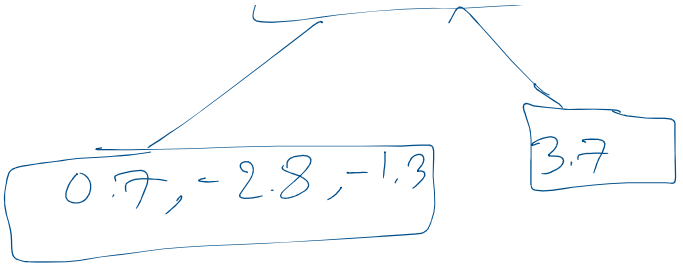
Gain = $SS_L + SS_R - SS_{Root}$

= 3.85 + 13.69 − 0.02

= 17.52

Steps - 2

Iq < 8.25



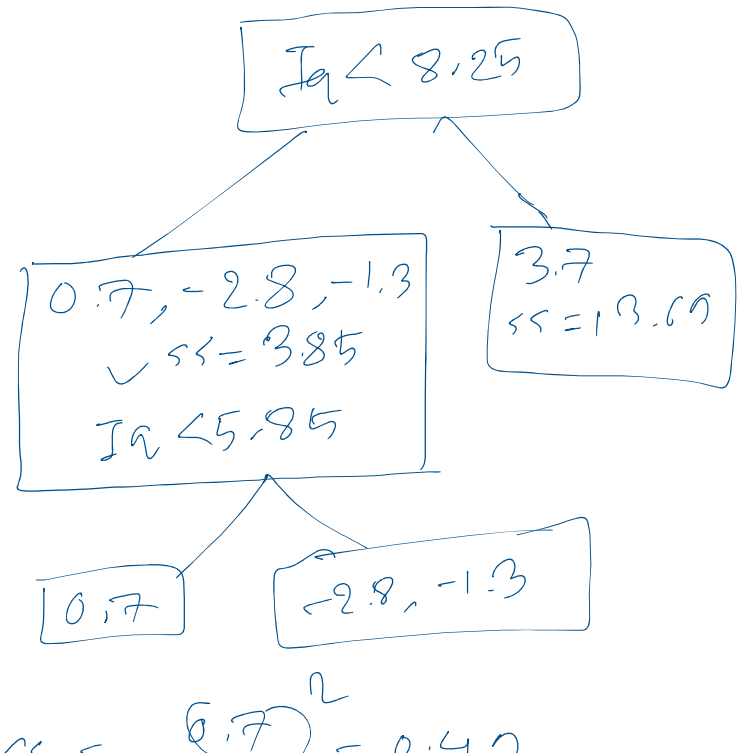
Iq	resi
5.0	0.7
6.7	-2.8
7.5	-1.3
9.0	3.7

L K R

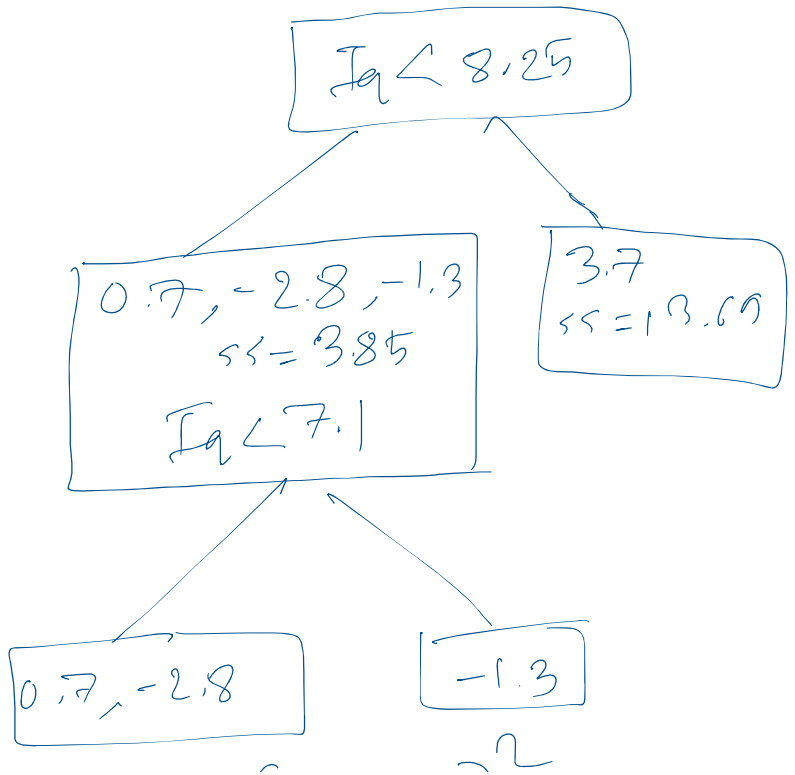
5.0] 5.85
6.7]
7.5] 7.1

Iq	resi
5.0	0.7
6.7	-2.8
7.5	-1.3

① sp 5.85



② sp 7.1



$(0.7)^2 - 0.49$

$SS_L = \frac{(0.7)^2}{1} = 0.49$

$SS_R = \frac{(-2.8-1.3)^2}{2+0} = \frac{(-4.1)^2}{2} = 8.40$

gain $= 0.49 + 8.40 - 3.85$

$= 5.04$ ✓

0.7, -2.8 | -1.3

$SS_2 = \frac{(0.7-2.8)^2}{2+0} = 2.20$

$SS_R = \frac{(-1.3)^2}{1} = 1.69$

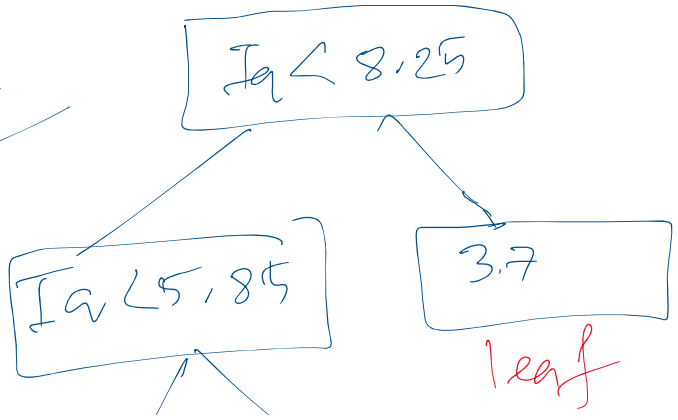
gain $= SS_2 + SS_R - SS_{parent}$

$= 2.20 + 1.69 - 3.85$

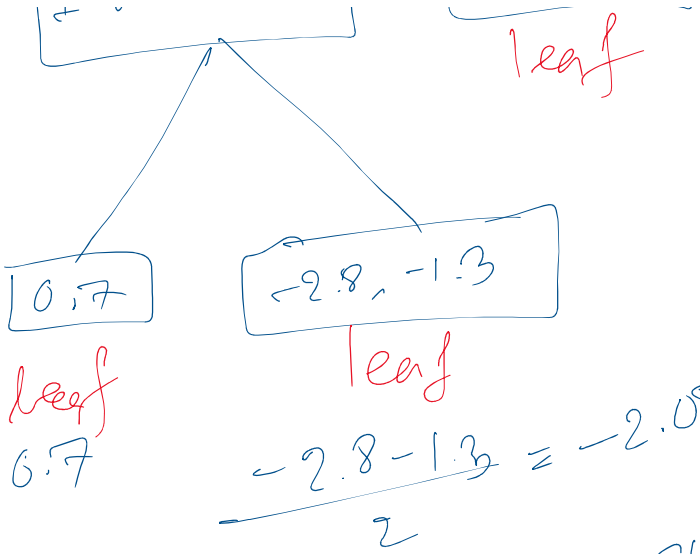
$= 0.04$ ↑

stage - 2 DT

depth : 2



output = $\frac{\text{sum of residuals}}{\text{Number of residuals} +}$



$$\frac{-2.8-1.3}{2} = -2.05$$

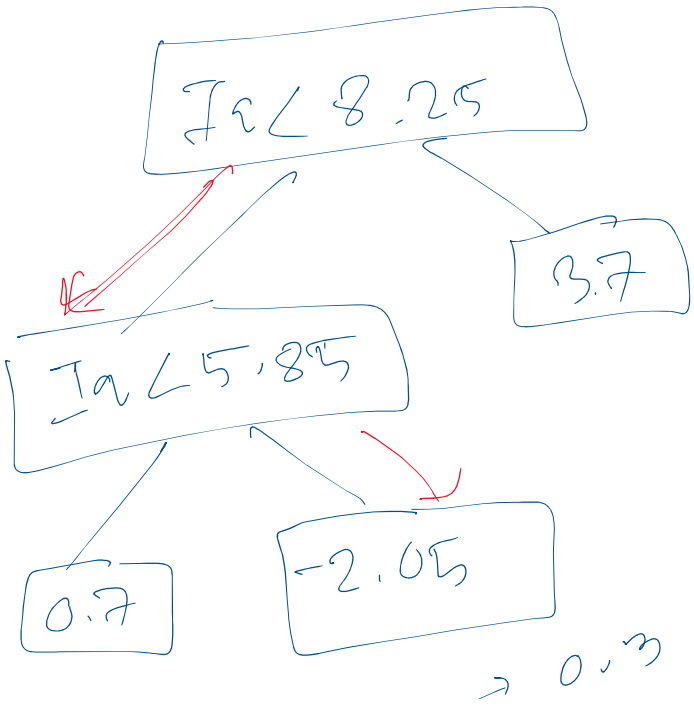
Stage - 2

DT

salary - model 2

Ia	salary	model 1	res 1	model 2	res 2
6.7	4.5	7.3	-2.8	6.68	-2.18
9.0	11.0	7.3	3.7	8.41	2.59
7.5	6.0	7.3	-1.3	6.68	-0.68
5.0	8.0	7.3	0.7	7.57	0.4

res → 0



$n_{obs} \to 0$

0.7

0.3

model2 = model1 + lr * (model2_

Eq = p = 7.3 + 0.3 x (0.7)

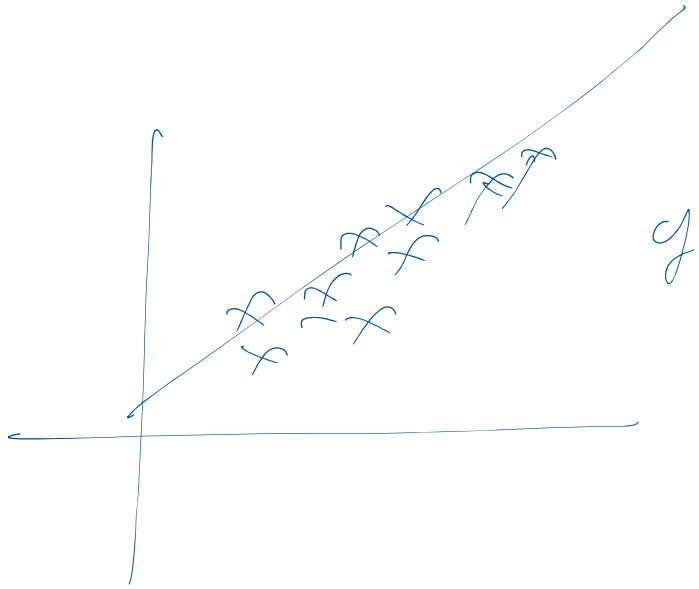
= 7.51

est_ lm(

Additive Model

$X \to Y$

$DT-P)$



$y = mx$

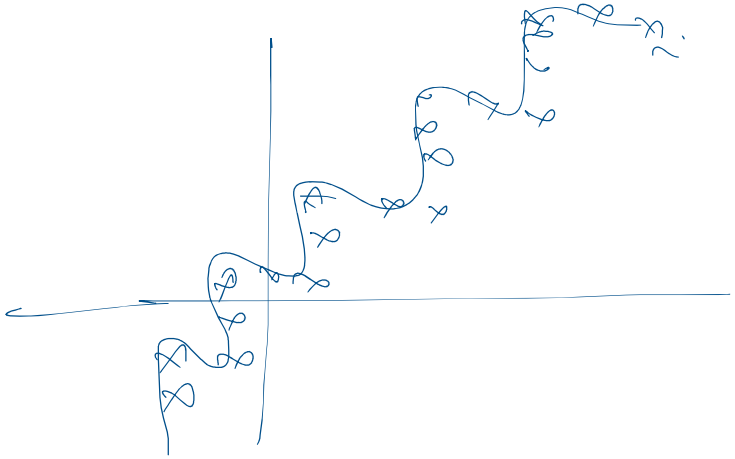
$f(m)$

$y \longrightarrow f(x)$

$y = f(m)$

$y = f(m_1) \pm f(m_2)$

$f \ldots$



$y = x + \cos x$