Lecture 01

Solved Example 01

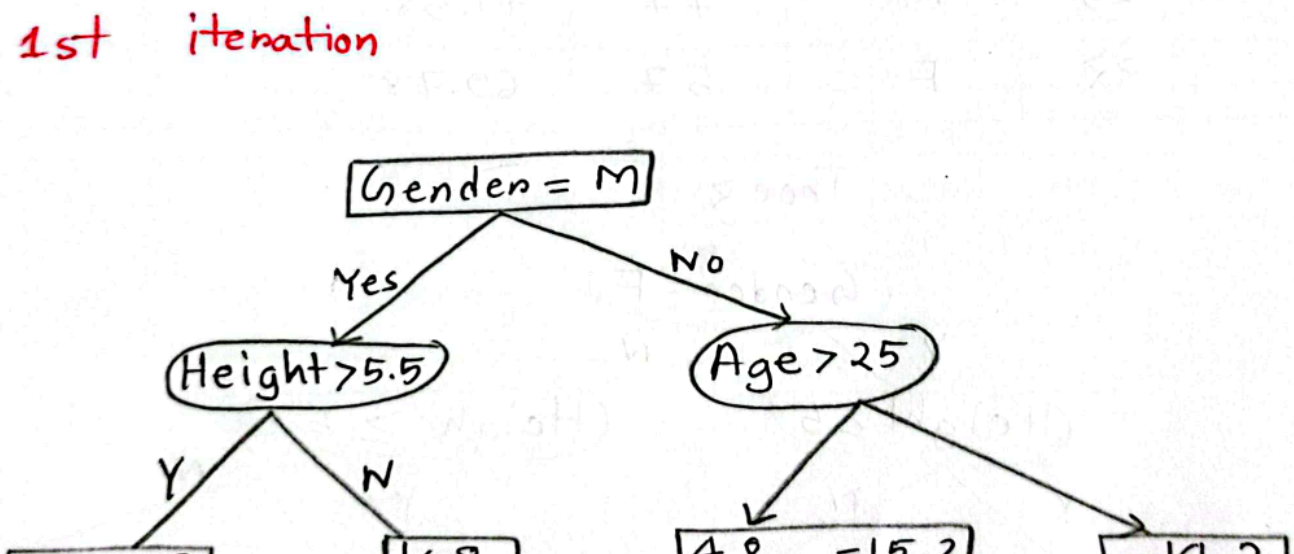
① $-\sum_{i=1}^{m} (\text{Observed}_i - \text{Predicted}_i) = 0$

$-[(88-P) + (76-P) + (56-P) + (73-P) + (77-P) + (57-P)] = 0$

$P = \frac{427}{6} = 71.2$

Height	Age	Genden	Weight Observed	Initial Pred. Value $f_0(x)$	Residuals P_i Observed $- f_0(x)$
5.4	28	M	88	71.2	16.8
5.2	26	F	76	71.2	4.8
5	28	F	56	71.2	-15.2
5.6	25	M	73	71.2	1.8
6	25	M	77	71.2	5.8
4	22	F	57	71.2	-14.2

Tree 1st iteration



$\frac{1.8 + 5.8}{2} = 3.8$

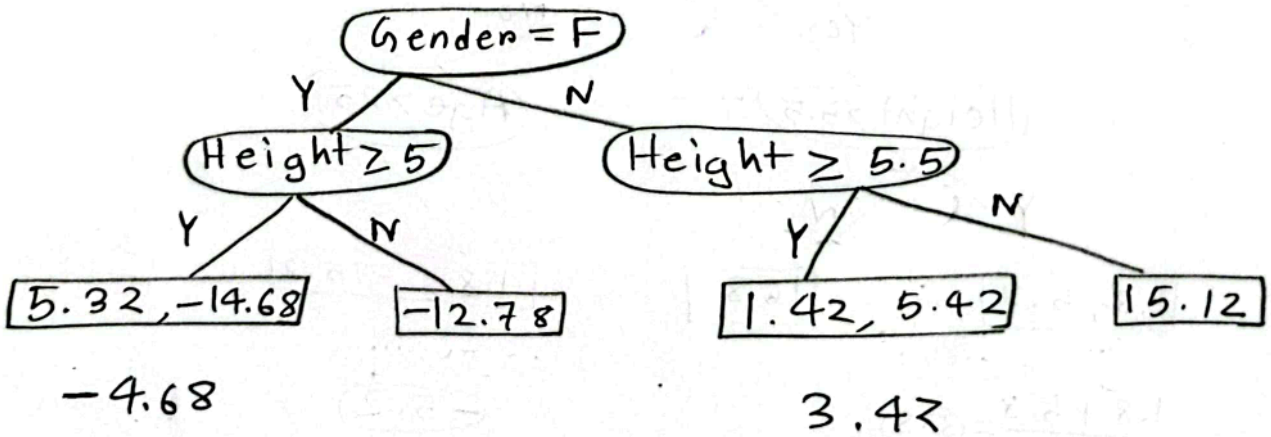
-5.2

weight	initial Pred	Residuals	Residuals from Tree	Prediction $F_1(x)$ = initial + α × Residual from Tree
88	71.2	16.8	16.8	71.2 + 0.1 × 16.8 = 72.88
76	71.2	4.8	−5.2	71.2 + 0.1 × (−5.2) = 70.68
56	71.2	−15.2	−5.2	71.2 + 0.1 × (−5.2) = 70.68
73	71.2	1.8	3.8	71.2 + 0.1 × (3.8) = 71.58
77	71.2	5.8	3.8	71.2 + 0.1 × (3.8) = 71.58
57	71.2	−14.2	−14.2	71.2 + 0.1 × (−14.2) = 69.78

2nd iteration:

Height	Age	Genden	weight	$F_1(x)$	Residuals = observed − $F_1(x)$
5.4	28	M	88	72.88	15.12
5.2	26	F	76	70.68	5.32
5	28	F	56	70.68	−14.68
5.6	25	M	73	71.58	1.42
6	25	M	77	71.58	5.42
4	22	F	57	69.78	−12.78

Tree 2



−4.68

3.42

Height	Prediction $F_1(x)$	Residual from Tree	$F_2(x) = F_1(x) + \alpha *$ Res. from Tree
88	72.88	15.12	$72.88 + 0.1 \times 15.12 = 74.4$
76	70.68	−4.68	$70.68 + 0.1 \times (-4.68) = 70.2$
56	70.68	−4.68	$70.68 + 0.1 * (-4.68) = 70.2$
73	71.58	3.42	$71.58 + 0.1 \times (3.42) = 71.9$
77	71.58	3.42	$71.58 + 0.1 \times (3.42) = 71.9$
57	69.78	−12.78	$69.78 + 0.1 \times (-12.78) = 68.5$

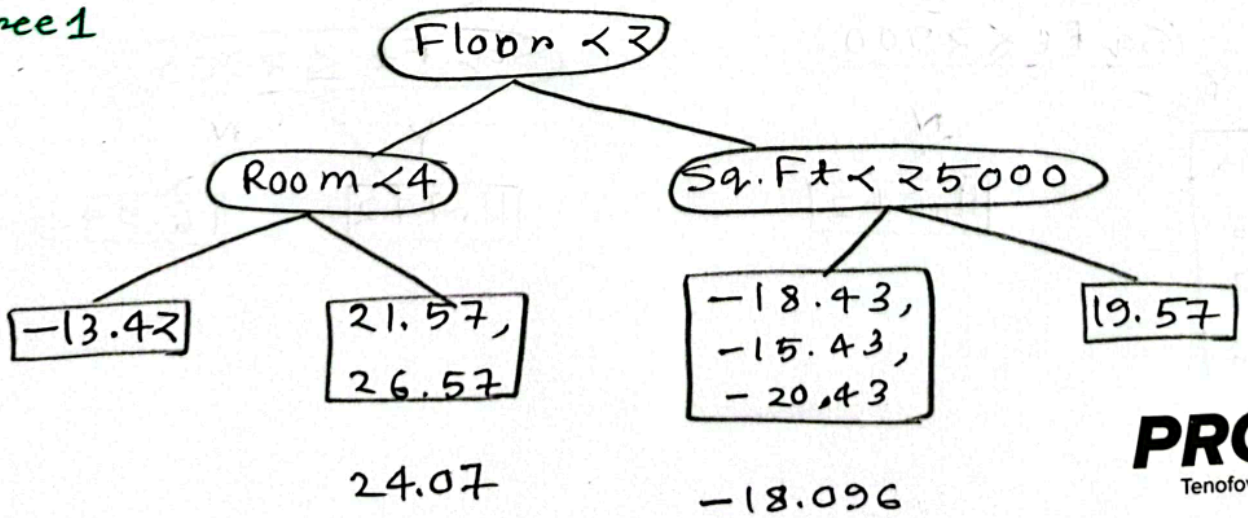
Solved Example 02

$$-[(130-P) + (90-P) + (93-P) + (128-P) + (95-P) + (88-P) + (135-P)$$

$$\therefore P = 108.43$$

Sq.Ft	Floor	Rooms No	Home Price	Initial Pred	Residuals
3300	1	4	130	108.43	21.57
2200	2	3	90	108.43	−18.43
2400	3	2	93	108.43	−15.43
2900	2	5	128	108.43	19.57
1800	1	3	95	108.43	−13.42
1700	3	4	88	108.43	−20.43
3400	1	5	135	108.43	26.57

Tree 1



Ada Boost Ensemble Learning

if CGPA ≥ 9 = Yes

CGPA	Pred Job offer	Actual Job offer	Weight
≥9	Yes	Yes	1/6
<9	No	Yes	1/6
≥9	Yes	NO	1/6
<9	No	NO	1/6
≥9	Yes	Yes	1/6
≥9	Yes	Yes	1/6

~~Iteration 1~~

S1: initial weight = 1/6

S2: total error = $\frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$

S3: weight $\alpha = \frac{1}{2} \ln\left(\frac{1 - \text{total error}}{\text{total error}}\right) = \frac{1}{2} \ln\left(\frac{1 - 1/3}{1/3}\right)$

$= 0.347$

S4: Wrong prediction = $\text{old}_w \cdot e^{\alpha} = \frac{1}{6} \times e^{0.347}$

$= 0.2358$

Correct prediction = $\frac{1}{6} \times e^{-0.347} = 0.1178$

Now,

2 x Wrong + 4 x Correct = $2 \times 0.2358 + 4 \times 0.1178$

$= 0.9428$

Normalization,

$N_{error} = \frac{0.2358}{0.9428} = 0.2501$

$N_{correct} = \frac{0.1178}{0.9428} = 0.1249$

Now total = $(2 \times 0.2501 + 0.1249 \times 4) = 0.999 \simeq 1$

Decision Stump for Interactiveness

Interactiveness	Pred Job offer	Actual Job offer	weight	N w
Yes	Yes	Yes	0.1249	0.
No	No	Yes	0.2501	0.
No	No	No	0.1249	0.
No	No	No	0.2501	
Yes	Yes	Yes	0.1249	0.
Yes	Yes	Yes	0.1249	0.

~~Iteration 2~~

S1 : initial weight = 0.1249

S2 : total error = $1 \times 0.2501 = 0.2501$

S3 : $\alpha = \frac{1}{2} \ln\left(\frac{1-0.2501}{0.2501}\right) = 0.549$

S4: Wrong pred = $0.2501 \times e^{0.549} = 0.433$

Correct pred = $0.1249 \times e^{-0.549} = 0.0721$

Here,

$$1 \times 0.4330 + 4 \times 0.0721 + 0.2501\, e^{-0.5490} = 0.8659$$

Normalize:

$$N_{error} = \frac{0.4330}{0.8659} = 0.5$$

$$N_{connect} = \frac{0.0721}{0.8659} = 0.0832$$

$$\frac{0.2501\, e^{-0.5490}}{0.8659} = 0.1668$$

Interactiveness	Pred Job offer	Actual Job offer	weight
Yes	Yes	Yes	0.0832
NO	NO	Yes	0.5001
NO	NO	NO	0.1667
NO	NO	NO	0.0832
Yes	Yes	Yes	0.0832
Yes	Yes	Yes	0.0832

ecision stump for Practical knowledge

Practical knowledge	Pred Job offer	Actual Job offer	Weight
Good	Yes	Yes	0.0832
Good	Yes	Yes	0.5001
Average	No	NO	0.1667
Average	No	NO	0.0832
Good	Yes	Yes	0.0832
Good	Yes	Yes	0.0832

S1: initial weight = 0.0832

S2: Total error = 0

so no need to change the weight.

iv. Decision stump for Communication skill

Communication skill	Pred Job offer	Actual Job offer	Weight
Good	Yes	Yes	0.083
Moderate	No	Yes	0.500
Moderate	No	No	0.166
Good	Yes	No	0.083
Moderate	No	Yes	0.083
Moderate	No	Yes	0.083

S1: initial weight

S2: Total error = $1 \times 0.5001 + 3 \times 0.0832 = 0.7497$

$$\alpha = \frac{1}{2} \ln\left(\frac{1-0.7497}{0.7497}\right) = -0.5485$$

$$Z_{cs} = 0.0832 \times e^{-(-0.5485)} + 0.5001 \times e^{-0.5485} + 0.1667 \times e^{-(-0.5485)} + 3 \times 0.0832 \times e^{-0.5485}$$

$$= 0.866$$

Update weights

Com. skill	Pred Job offer	Actual Job offer	Weight
Good	Yes	Yes	$\frac{0.0832 \times e^{-(-0.5458)}}{0.866} = 0.1663$
Moderate	No	Yes	$\frac{0.5001 \times e^{-0.5458}}{0.866} = 0.3337$
Moderate	No	No	$\frac{0.1667 \times e^{-(-0.5458)}}{0.866} = 0.3331$
Good	Yes	No	$\frac{0.0832 \times e^{-0.5458}}{0.866} = 0.0555$
Moderate	No	Yes	0.0555
Moderate	No	Yes	0.0555

$\alpha_{CGPA} = 0.347$	$\alpha_{internac} = 0.549$	$\alpha_{Com\,skill} = -0.5485$	Weighted Avg	Final Pred
Yes	Yes	Yes	0.347	Y
No	No	No	0	N
Yes	No	No	0.347	Y
No	No	Yes	−0.5485	N
Yes	Yes	No	0.896	Y
Yes	Yes	No	0.896	Y

Solved Problem 01 (Regression) (XGBoost)

Drug Dosage	Drug effect	$y_i^{(0)}$	$r_i^{(1)}$
12	−10	0.5	−10.5
22	7	0.5	6.5
28	8	0.5	7.5
32	−7	0.5	−7.5

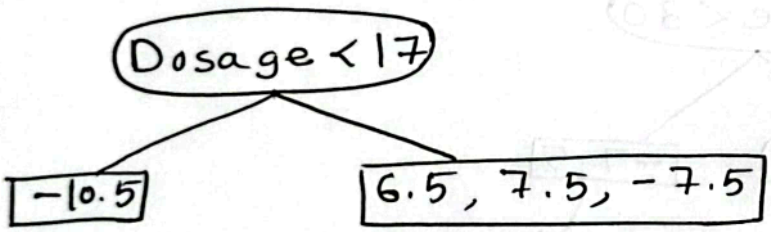
① Initialize $y_i = 0.5$ [ডেটা নিলে] and calculate residuals.

② similarity score $= \frac{(-10.5 + 6.5 + 7.5 + -7.5)^2}{4+\lambda}$ $(\lambda = 1)$

$= 3.2$

③ A

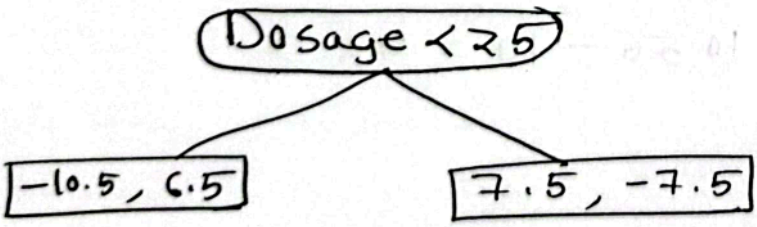
sc ↑ split better.



$sc = \frac{(6.5)^2}{3+1} = 10.56$

$sc = 55.13$

B



$sc = 0$

$sc = \frac{(-10.5+6.5)^2}{2+1}$

$= 5.33$

④ Calculate gain

✓ For Tree A,

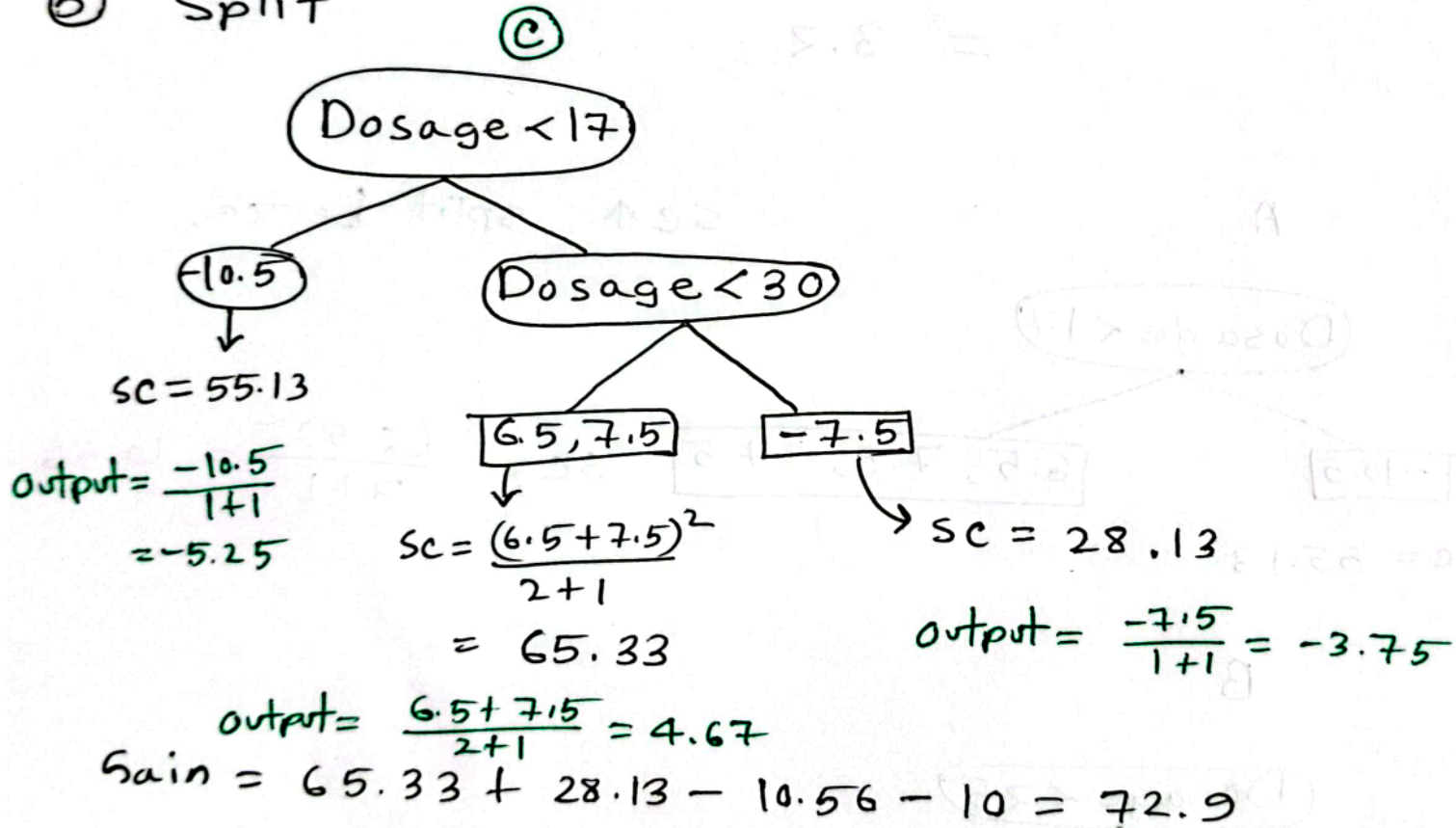
gain = Left sc + Right sc − sc before split − γ

= (55.13 + 10.56 − 3.2 − 10) = 52.49

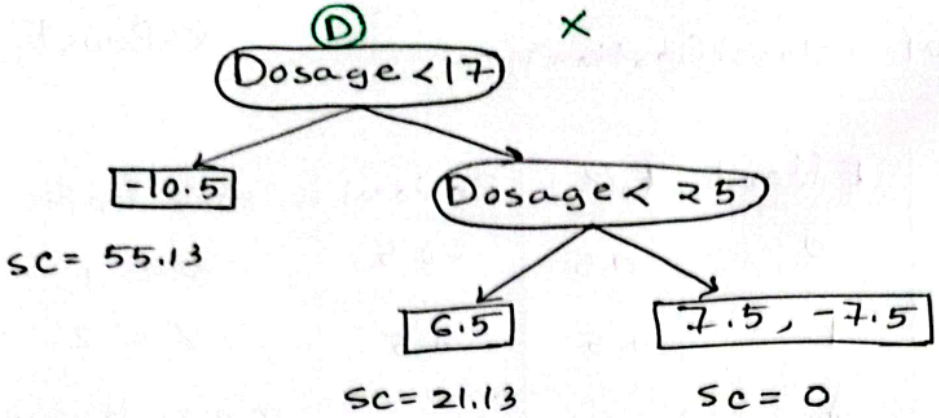
✗ For Tree B,

gain = 5.33 + 0 − 3.2 − 10 = −7.87

⑤ Split



Gain = 65.33 + 28.13 − 10.56 − 10 = 72.9



SC = 55.13

SC = 21.13

SC = 0

gain = 21.13 + 0 − 10.56 − 10 = 0.57

⑥ output value $(w_j) = \dfrac{\sum p_i^{(1)}}{|\text{residual}| + \lambda}$

⑦ calculate new predictions

$$\hat{y}_i^{(1)} = \hat{y}_i^{(0)} + \eta w_j$$

$$\hat{y}_1^{(1)} = 0.5 + 0.3 \times (-5.25) = -1.075$$

$$\hat{y}_2^{(1)} = 0.5 + 0.3 \times (4.67) = 1.901$$

$$\hat{y}_3^{(1)} = 0.5 + 0.3 \times (4.67) = 1.901$$

$$\hat{y}_4^{(1)} = 0.5 + 0.3 \times (-3.75) = -0.625$$

⑧ New Residuals

$$p_1^{(2)} = -10 - (-1.075) = -8.925$$

$$p_2^{(2)} = 7 - (1.901) = 5.099$$

$$p_3^{(2)} = 8 - (1.901) = 6.099$$

$$p_4^{(2)} = -7 - (-0.625) = -6.375$$

(Repeat 2 to 8)

The Residuals continu
to decrease and predic
y gets closer to the

XgBoost

Solved Problem 01 (classification)

step 1 and 2

Dosage	Gender	Effect	$F_0(x)$	Residual
2	M	0	0.5	-0.5
8	F	1	0.5	0.5
12	M	1	0.5	0.5
18	F	0	0.5	-0.5

max depth = 2

$\alpha = 1$

$\gamma = 2$

min child weight = 0

reg_lambda $\lambda = 0$

base_score = 0.5

n_estimator = 2

step 3

Root Gain = 1 + 0.33 − 0 = 1.33

Dosage < 5

$Sc = \frac{0^2}{[0.5(1-0.5) + 0.5(1-0.5) + 0.5(1-0.5) + 0.5(1-0.5)] + 0} = 0$

−0.5 | 0.5, 0.5, −0.5

$Sc = \frac{(-0.5)^2}{0.5(1-0.5)+0} = 1$

$Sc = \frac{(0.5)^2}{[0.5(1-0.5) + 0.5(1-0.5) + 0.5(1-0.5)]+0} = 0.33$

Now

Dosage < 10

R_Sc = 0

Gain = 0

−0.5, 0.5 | 0.5, −0.5

Sc = 0 | Sc = 0

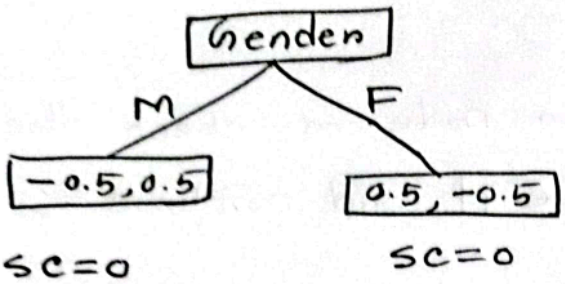
Dosage < 15

R_Sc = ~~1.33~~ 0

Gain = 1.33

−0.5, 0.5, 0.5 | −0.5

Sc = 0.33 | Sc = 1



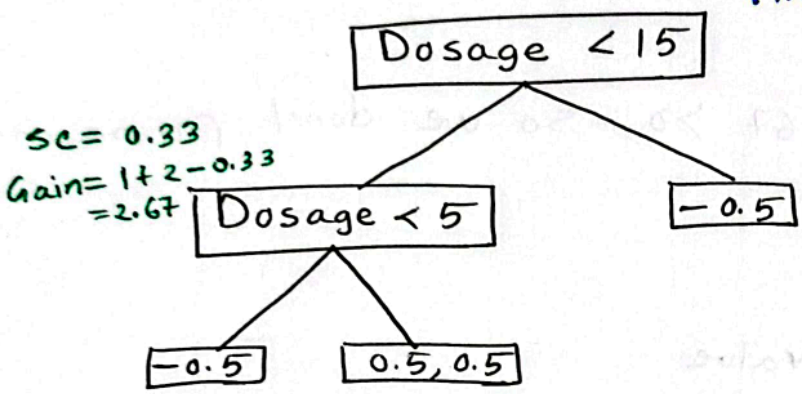
R_sc = 0

Gain = 0

sc = 0 sc = 0

we chose <15, max_depth = 2.

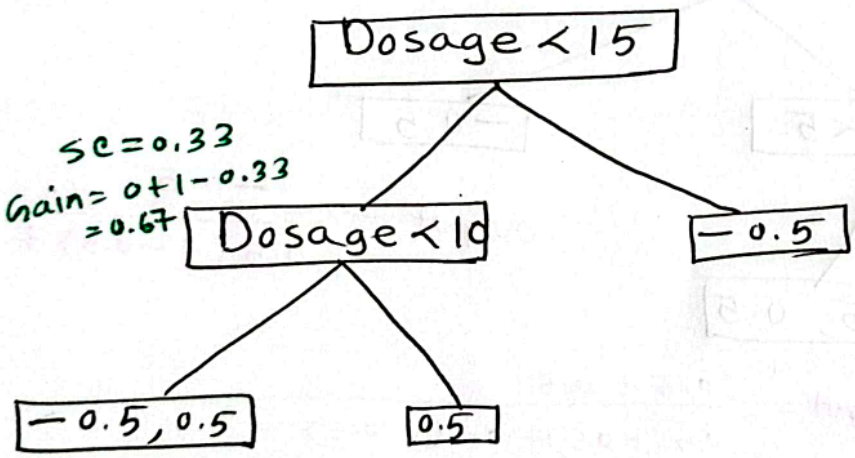
Tree 1



sc = 1

$$sc = \frac{(0.5+0.5)^2}{[0.5\times(1-0.5)+0.5\times(1-0.5)]+0} = 2$$

Tree 2 ⊗



sc = 0 sc = 1

From the fig based on gain we choose Dosage < 5 Tree 1 and constructed a fully grown tree.

If (Gain − γ) < 0 :

Remove that particular node and check the same for the node just above it and continue to pruning.

Else :

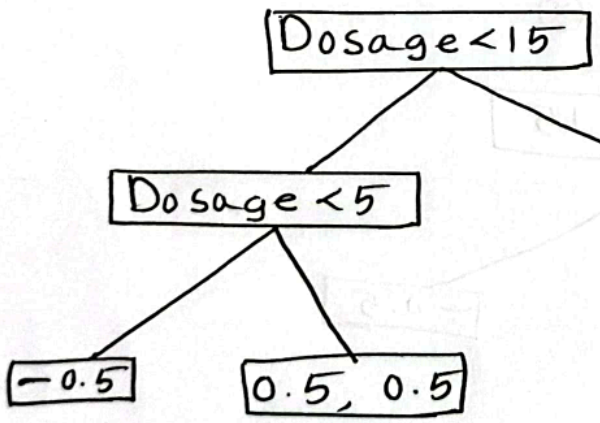
Do not remove that node and stop pruning.

Here Gain = 2.67

so, (2.67 − 2) = 0.67 > 0, so we dont prune and leave as it is.

Calculate output value

$$\text{output} = \frac{\sum \text{Residual}}{\left[\sum \text{Prev. Prob.} * (1 - \text{Pre. Prob.}) + \lambda\right]}$$



output = −2

output = $\frac{0.5 + 0.5}{0.5(1-0.5) + 0.5(1-0.5)}$

= 2

output = $\frac{-0.5}{0.5(1-0.5) + 0}$ = −2

$$F_0(x) = h_0(x) = \log\left(\frac{p}{1-p}\right) = \log\left(\frac{0.5}{1-0.5}\right) = 0$$

(p → initial pred.)

$$F_1(x) = \text{sigmoid}\left(h_0(x) + \alpha * h_1(x)\right) \quad \boxed{\alpha = 1}$$

($h_1(x)$ → output)

$$= \text{sigmoid}(0 + 1 * -2) = \text{sigmoid}(-2)$$

$$= \frac{1}{1+e^{-(-2)}} = 0.119$$

Again, $\text{sigmoid}(2) = \frac{1}{1+e^{-2}} = 0.88$

Dosage	Gender	Effect	$h_0(x)$	$F_1(x)$
2	M	0	0	0.11
8	F	1	0	0.88
12	M	1	0	0.88
18	F	0	0	0.11

Repeat step 2 and 3 until required number of models are build.

Feature Extraction (PCA)

① calculate mean of Every feature.

② calculate co-variance matrix

③ calculate Eigen values and Eigen vectons

④ choose Principal components

⑤ Project Data onto Principal component.

1.

x_1	x_2
4	11
8	4
13	5
7	14
Mean = 8	8.5

$N = 4$

2. $$cov(x_i, x_j) = \frac{1}{N-1} \sum_{k=1}^{N} (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)$$

$$cov(x_1, x_1) = \frac{1}{4-1}\left[(4-8)^2 + (8-8)^2 + (13-8)^2 + (7-\right.$$

$$= 14$$

$$cov(x_1, x_2) = \frac{1}{4-1}\left[(4-8)(11-8.5) + (8-8)(4-8.5) + (13-8)(5-8.5) + (7-8)(14-8.5)\right]$$

$$= -11$$

$$\text{cov}(x_2, x_1) = -11$$

$$\text{cov}(x_2, x_2) = \frac{1}{4-1}\left[(11-8.5)^2 + (4-8.5)^2 + (5-8.5)^2 + (14-8.5)^2\right]$$

$$= 23$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

3. Eigen values and Eigen vectors

Set up the equation $\det(S - \lambda I) = 0$

$$\begin{vmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (14-\lambda)(23-\lambda) + 121 = 0$$

$$\Rightarrow 322 - 14\lambda - 23\lambda + \lambda^2 + 121 = 0$$

$$\Rightarrow \lambda^2 - 37\lambda + 201 = 0$$

$$\lambda_1 = 30.38$$

$$\lambda_2 = 6.615$$

Eigen vecton

$$U = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$(S - \lambda_1 I)U = 0$$

$$\begin{bmatrix} 14-\lambda_1 & -11 \\ -11 & 23-\lambda_1 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow (14-\lambda_1)U_1 - 11U_2 = 0 \quad \text{---------} \quad (1)$$

$$-11U_1 + (23-\lambda_1)U_2 = 0 \quad \text{---------} \quad (11)$$

From (1) $\Rightarrow$

$$\frac{(14-\lambda_1)U_1}{(14-\lambda_1)11} - \frac{11U_2}{(14-\lambda_1)11} = 0$$

$$\Rightarrow \frac{U_1}{11} - \frac{U_2}{(14-\lambda_1)} = 0$$

Let,

$$\frac{U_1}{11} = \frac{U_2}{(14-\lambda_1)} = t \text{ (any real number)}$$

taking $t = 1$,

$$\frac{U_1}{11} = \frac{U_2}{14-\lambda_1} = 1$$

$\therefore U_1 = 11 \quad \& \quad U_2 = 14-\lambda_1 = -16.38$

$$U = \begin{vmatrix} 11 \\ -16.38 \end{vmatrix} \quad \text{[eigen vector]}$$

$$U = \sqrt{(11)^2 + (-16.38)^2} = 19.73$$

$$e_1 = \left| \begin{matrix} \frac{11}{19.73} \\ \frac{-16.38}{19.73} \end{matrix} \right| = \left| \begin{matrix} 0.5575 \\ -0.8302 \end{matrix} \right| \begin{matrix} \rightarrow x \\ \rightarrow y \end{matrix}$$

7. PCA

x_1	x_2	PCA
4	11	−4.3055
8	4	3.7359
13	5	5.6932
$\frac{7}{8}$	$\frac{14}{8.5}$	−5.1236

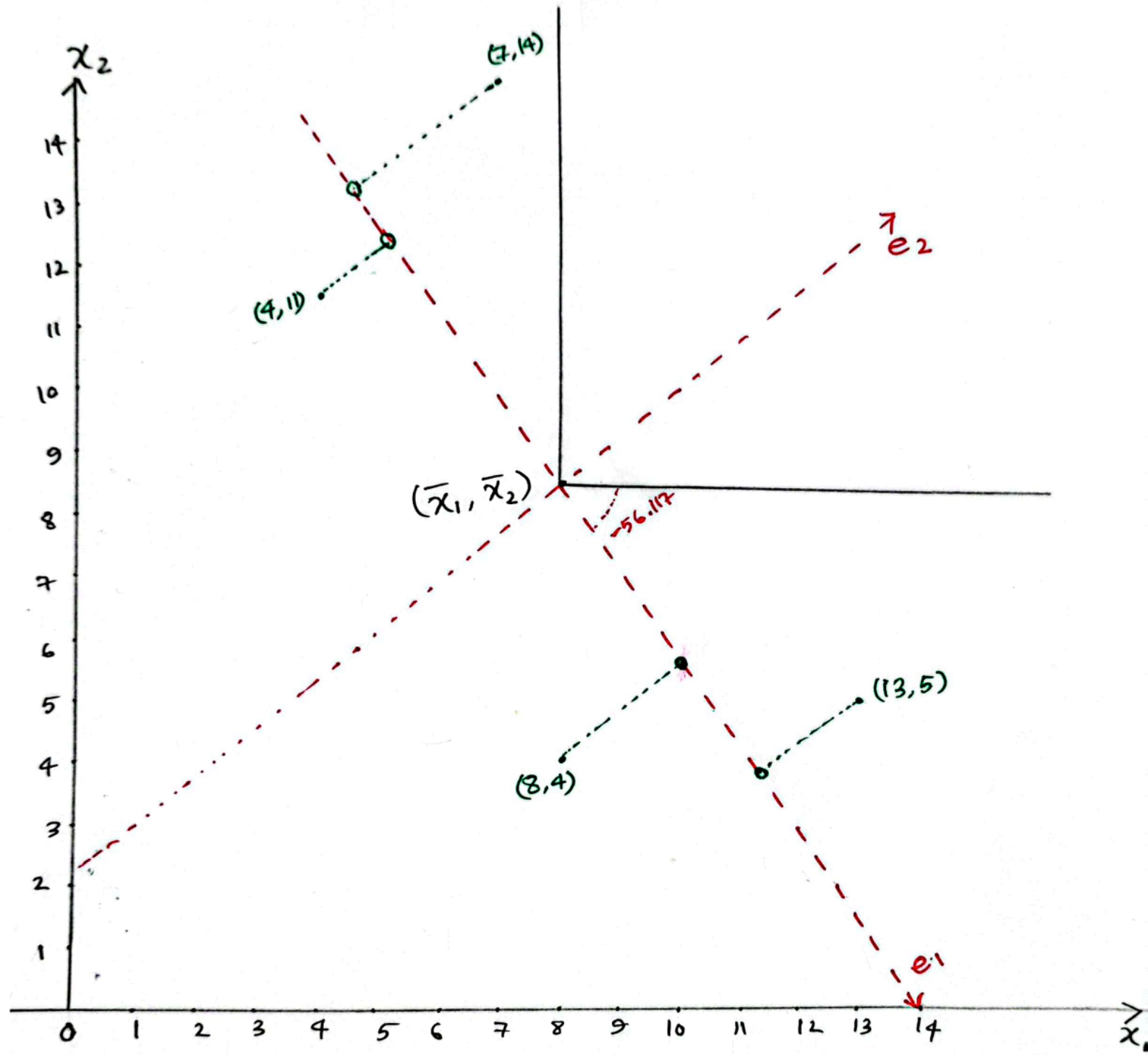
$$\theta = \tan^{-1} \frac{y}{x}$$

$$e_1^T \begin{bmatrix} x_1 - \bar{x}_1 \\ x_2 - \bar{x}_2 \end{bmatrix}$$

$$PCA_1 = \begin{bmatrix} 0.5575 & -0.8302 \end{bmatrix} \begin{bmatrix} 4-8 \\ 11-8.5 \end{bmatrix} = -4.3055$$

. Project Data onto Principal component.

$$\theta = \tan^{-1}\left(\frac{-0.8302}{0.5575}\right) = -56.117$$


x_2
(7,14)
(4,11)
e_2
$(\bar{x}_1, \bar{x}_2)$
56.117
(13,5)
(8,4)
e_1
x_1
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
1 2 3 4 5 6 7 8 9 10 11 12 13 14