Gradient Boosting (classification)

chocolate

	Choc Pref. (f_1)	Age (f_2)	Pref. color (f_3)	Sugar Pref (target)	
1	Y	12	B	Y (1)	→ 0.3
2	Y	87	G	Y (2)	→ 0.3
3	N	44	B	N	→ −0.7
4	Y	19	R	N	→ −0.7
5	N	32	G	Y (3)	→ 0.3
6	N	14	B	Y (4)	→ 0.3

R_1

P_1 → M1

① $\log(\text{odds}) = \log_e\left(\frac{4}{2}\right) = \log_e 2 = 0.7$

(Rounded off)

② Probability of Sugar Preference =

$$\frac{e^{0.7}}{1+e^{0.7}} = 0.7 \approx 0.66$$

③ Residual → Actual − Predicted

$m_2 \longrightarrow$ Pref. color = Red

yes / no

Decision Tree

-0.7 → Residuals

-3.3

Age > 37

yes / no

Residuals

0.3, -0.7

-1

0.3, 0.3, 0.3

1.4

Transformation

Σ Residuals

Σ [Previous Probability$_i$ * (1 - PP$_i$)]

Leaf-1

$$\frac{-0.7}{0.7 * (1-0.7)}$$

$$= -3.3$$

Leaf-2

$$\frac{0.3 - 0.7}{2\,(0.7 * (1-0.7))} = -1$$

Leaf-3

$$\frac{0.3 + 0.3 + 0.3}{3\,(0.7 + (1-0.7))} = 1.4$$

Learning Rate

(m1 & m2) m1 0.8 m2

new Prediction → 0.7 + (Y * 1.4)

⇒ 1.8

0.7 + 0.8 * -1 = -0.1

Probability → $\frac{e^{1.8}}{1+e^{1.8}} = 0.9$

→ $\frac{e^{-0.1}}{1+e^{-0.1}} = 0.5$ → m3 (target (R2))

(f1) (f2) (f3) target

Choc Pref.	Age	Pref. color	Sugar	P1	R1	P2	R2
Y	12	B	Y(1)	0.7	0.3	0.9	0.1
Y	87	G	Y(1)	0.7	0.3	0.5	0.5
N	44	B	N(0)	0.7	-0.7	0.5	-0.5
Y	19	R	N(0)	0.7	-0.7	0.1	-0.1
N	32	G	Y(1)	0.7	0.3	0.9	0.1
N	14	B	Y(1)	0.7	0.3	0.9	0.1
Y	25	G	1				

(Y)

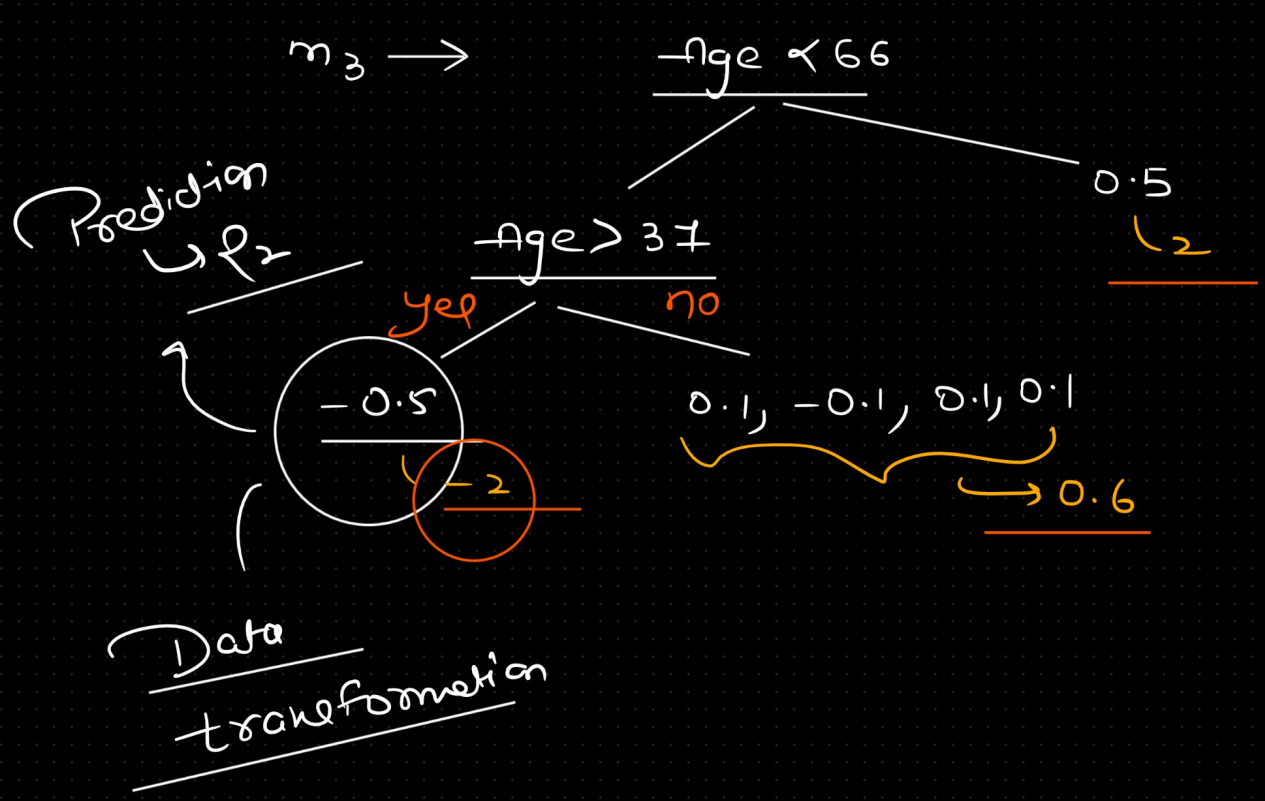
new test data → to combine Prediction by (m1, m2 & m3)

0.7 + 0.8 * 1.4 + 0.8 * 0.6

m1 η ⇒ 2.3

Probability → $\frac{e^{2.3}}{e^{2.3}+1}$ → 0.9

0.5



Data transformation