Logistic Regression

Hrs (studied)	outcome (y)
1	0
2	0
3	0
4	1
5	1
6	1

$$P(y=1|x) = 1 / 1+e^{-(a_0+a_1x)}$$

$a_1 = 0.7212$, $a_0 = -3.0254$.

[using gradient descent tech]

$$P(y=1|x) = 1/1+e^{-(-3.0254+0.7212x)}$$

when $x=1$, $P(y=1|1)$

$$= 1/1+e^{-(-3.0254+0.721212.1)}$$

$$\approx 0.090$$

(Fail)

$x=2,$

$$P(y=1|2) = \frac{1}{1+e^{-(-3.0254+0.7212121\times2)}}$$

$= 0.170$ (Fail)

$x=3,$ -0.247 (Fail)

$x=4,$ $P(y=1|4) = 0.247$ (Fail)

$x=5$ $P(y=1|5) = 0.641$ (~~Fail~~ Pass)

$x=6,$ $P(y=1|6) = 0.780$ (Pass)

$P(y=1|x) \geq 0.5 \Rightarrow P$

$P(y=1|x) < 0.5 = F$

prediction

Polynomial Regression

x	y
1	2
2	4
3	6

$\Rightarrow$

$$y = \underbrace{a_0 + a_1 x + a_2}_{\text{co-efficients}} x^2$$

Note: $A^{-1} = \frac{1}{|A|} * adj(A)$

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $|A| = ad - bc$

adj(A) = swap, a & d, negate b & c

$adj(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

$$A^{-1} = \frac{1}{ad-bc} * \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

x	y
1	2
2	3
3	5
4	4
5	6

2-nd degree polynomial

$$y = a_0 + a_1 x + a_2 x^2$$

create design matrix

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ 1 & x_5 & x_5^2 \end{bmatrix}$$

dataset = {$x = [1, 2, 3, 4, 5]$}

$$X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix}$$

Target vector $y = \begin{bmatrix} 2 \\ 3 \\ 5 \\ 4 \\ 6 \end{bmatrix}$

solve for coefficient

$$a = (X^T X)^{-1} X^T y$$

1) $X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 9 & 16 & 25 \end{bmatrix} * \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix}$

$= \begin{bmatrix} 5 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{bmatrix}$

2) $X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 9 & 16 & 25 \end{bmatrix} * \begin{bmatrix} 2 \\ 3 \\ 5 \\ 4 \\ 6 \end{bmatrix}$

$= \cancel{\begin{bmatrix} 20 \\ 70 \\ 278 \end{bmatrix}} \begin{bmatrix} 20 \\ 50 \\ 240 \end{bmatrix}$

$$(X^T X)^{-1} = \begin{bmatrix} 11.5 & -3.25 & 0.5 \\ -3.025 & 1.25 & -0.25 \\ 0.5 & -0.25 & 0.0625 \end{bmatrix}$$

$$(X^T X^{-1}) X^T Y \Rightarrow \frac{1}{} * \begin{bmatrix} 20 \\ 50 \\ 240 \end{bmatrix} =$$

$$\theta = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 1.3285714.3 \\ -0.07142857 \end{bmatrix}$$

$x = 5$

$$\Longrightarrow y = 0.8 + 1.3857143 * 5 - 0.0714285 7 * 5^2$$

$$y = 5.6571 4286$$

$$\boxed{y = 5.66}$$

Decision Tree

Study hrs	Attendance	P/F
High →	High →	Pass
Low →	High →	Pass
Medium →	High →	Pass
Low →	Low →	Fail
Medium →	Low →	Fail

Steps

1) Calculate entropy for the dataset
2) Calculate information gain
3) Choose the best feature to split
4) Repeat until Tree is complete

Step 1 The dataset has 3-P, 2-F

Formula, Entropy $H = -\sum p_i \times \log_2(p_i)$

$P_{pass} = 3/5 = 0.6$

$P_{Fail} = 2/5 = 0.4$

$$H = -[0.6 \cdot \log_2(0.6) + 0.4 * \log_2(0.4)]$$

$$H = -[0.6 * -0.737 + 0.4 * -1.322]$$

$$= -[-0.4422 - 0.522]$$

$$= -(-0.971) = \underline{\underline{0.971}}$$

Datased Entropy = 0.971

2) calculate information gain

① Study Hours $-\sum p_i \log_2(p_i)$

Study(hrs)	P	F	Total	H
High	1	0	1	0
Medium	1	1	2	1
Low	1	1	2	1

weighted entropy $= \frac{1}{5}(0) + \frac{2}{5}(1) + \frac{2}{5}(1)$

$= 0.8$

Info gain = $H_{Datased} - H_{Feature}$

$= 0.971 - 0.8 = \underline{0.171}$

2) Attendance

Attendance	P	F	Total	H
High	3	0	3	0
Low	0	2	2	0

weighted entropy $= \left(\frac{3}{5}\right)0 + \left(\frac{2}{5}\right)0 = 0$

gain $= 0.971 - 0 = 0.971$

choose the best Feature

Highest Inf gain → Attendance

∴ root node of tree ⇒ Attendance

Final Dicision Tree

Attendance

High (Pass)

Low (Fail)

Study hrs = Low, Attendance = High → Pass

Study hrs = Medium, —"— = Low → Fail

Example 2

Weather	Play
Sunny	No
Sunny	No
Overcast	Yes
Rain	Yes
Rain	No

Classify weather play based on weather

1) Calculate Dataset Entropy

Yes $\Rightarrow$ 2 , No $\Rightarrow$ 3

$$H_{Dataset} = -P_{yes} * \log_2(P_{yes}) - P_{No} * \log_2(P_{No})$$

$$P_{yes} = 2/5 \quad , \quad P_{No} = 3/5$$

$$H_{Dataset} = -(2/5 * \log_2(2/5)) - (3/5 * \log_2(3/5))$$

$$= -(0.4 * -1.322) - (0.6 * -0.737)$$

$$= 0.971$$

Weather	Y	N	Total	Δ
Sunny	0	2	2	0
Overcast	1	0	1	0
Rain	1	1	2	1.0

weighted entropy

$$H_{weather} = (2/5)\,0 + (1/5)\,0 + (2/5)\,1$$

$$= \underline{\underline{0.4}}$$

Info gain $\quad Gain = H_{Dataset} - H_{weather}$

$$= 0.971 - 0.4$$

$$= \underline{\underline{0.571}}$$

weather

Sunny (No) | Overcast (Yes) | Rain (Mixed)