1.

Information gain =

entropy parent - weight average entropy (children)

entropy for node = $-p_1 \log_2(p_1) - p_2 \log_2(p_2)$

- Calculate entropy for parent node

entropy (parent) = $-0.5 \log_2(0.5) - 0.5 \log_2(0.5)$

$= -0.5(-1) - 0.5(-1)$

$= 1.0$

- Calculate entropy child node 1

credit score ≤ 650

entropy (CS ≤ 650)

$= -0 \log_2(0) - 1 \log_2(1)$

$= 0$

- calculate entropy child node 2

credit score > 650

entropy (CS > 650)

$= -1 \log_2(1) - 0 \log_2(0)$

$= 0.$

LHP

calculate average entropy

$$\text{wae} = (0.5 \times e(\leq 650)) + (0.5 \times e(> 650)) = 0.$$

calculate information gain.

$$\text{information gain} = e(\text{parent}) - \text{wae}(\text{children}) = 1.0 - 0 = 1.0.$$

2.

formula for

Variance reduction =

variance (parent) − weighted average

variance (children)

variance of a node

$$\text{variance} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

- calculate variance for parent

$$\bar{x} = \frac{720 + 650 + \dots 640}{8}$$

$$= 660$$

variance of parent

$$= \frac{1}{8}(720-660)^2 + (650-660)^2 + \dots$$

$$= \frac{1}{8} \times 38600 = 4200$$

- calculate variance age ≤ 35

$$\bar{x} = \frac{720 + 650 + \dots + 640}{5} = 648$$

variance (Age ≤ 35)

$$= \frac{1}{5}(720-648)^2 + \dots + (640-648)^2$$

$$= 1576$$

Calculate variance Age > 35

$$\bar{x} = \frac{750 + 780 + 710}{3} = 746,66$$

variance (Age > 35) =

$$\frac{1}{3}(750 - 746,7)^2 + \dots + (710 - 746,7)^2$$

$$= 822,2$$

Weight avg variance =

(0.625 × var (Age ≤ 35)) +
(0.375 × var (Age > 35))

$$= 1293,3$$

LHP

calculate variance reduction.

$VR = var(parent) - wav(children)$

$= 4200 - 1293,3 = 2906,7.$

3.

T2 Credit score =

=> T2 is high risk 100 %

(4/4 records are high risk)

T2 Age = 30

=> T2 is high risk 80 %

1 (Risk level = high |

Age ≤ 35, Credit score ≤ 650)

$= \frac{4}{4} = 1.0$

4

$\hat{y} = \theta_0 + \theta_1 \cdot Age$

$\theta_0 = 500, \theta_1 = 5$

$\alpha = 0.01$

To do :

calculate cost function

perform one iteration

formula for cost function

$$J(\theta_0, \theta_1) = \frac{1}{2n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2$$

$$\theta_0 \leftarrow \theta_0 - \alpha \cdot \frac{\partial J}{\partial \theta_0}$$

$$\theta_1 \leftarrow \theta_1 - \alpha \cdot \frac{\partial J}{\partial \theta_1}$$

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i) \cdot x_i$$

$$\hat{y}_i = \theta_0 + \theta_1 \cdot \text{Age}_i$$

$$= 500 + 5 \cdot \text{Age}_i$$

LHP

$$\hat{y}_i = 500 + 5 \cdot \text{Age}_i$$

$500 + 5 \cdot 35 = 675$

$500 + 5 \cdot 28 = 640$

...

$500 + 5 \cdot 33 = 665$

error $(\hat{y}_i - y_i)$

$675 - 720 = -45$

$640 - 650 = -10$

...

$665 - 640 = 25$

$$\sum (\hat{y}_i - y_i)^2 = (-45)^2 + (-10)^2 + \ldots + 25^2$$
$$= 7025$$

$$J(\theta_0, \theta_1) = \frac{1}{2 \times 8} \times 7025$$

$$= 439.06$$

LHP

- calculate gradient for θ_0

$$\sum (\hat{y}_i - y_i) = -45 + (-10) + \dots + 25$$
$$= -5$$

$$\frac{\partial J}{\partial \theta_0} = \frac{-5}{8} = -0,625$$

- calculate gradient for θ_1.

$$\sum (\hat{y}_i - y_i) \cdot x_i$$

$$= (-45 \times 85) + (-10 \cdot 28) + \dots + (25 \cdot 33)$$
$$= -1055$$

$$\frac{\partial J}{\partial \theta_1} = \frac{-1055}{8} = -131,875.$$

LHP

Update the parameters.

$$\alpha = 0.01$$

$$\theta_0 \leftarrow \theta_0 - \alpha \times \frac{\partial J}{\partial \theta_0}$$

$$\theta_0 = 500 - 0.01 \times (-0.625)$$
$$= 500 + 0.00625$$
$$= 500.00625$$

$$\theta_1 \leftarrow \theta_1 - \alpha \frac{\partial J}{\partial \theta_1}$$

$$\theta_1 = 5 - 0.01 \times (-131.875)$$
$$= 5 + 1.31875$$
$$= 6.31875.$$

The updated parameters is

$\theta_0 = 500.00625$

$\theta_1 = 6.31875.$

LHP