SML - Assignment-1

Prabhat Gupta
2021550

5. Every ~~datapoint~~ data point in the table.

A new data point has age = 37 & loan = 142

We can represent each data point as a tuple of (age, loan).

So, the new data point is represented as → (37, 142).

Let the new data pt. be represented by ~~a~~ $a_1 = (37, 142)$ &

all the elements in the table be represented by $b_1, b_2, \ldots, b_{11}$.

So, we need to find the distance b/w a_1 & all pts $b_1, b_2, \ldots, b_{11}$.

$$d(a_1, b_1) = \sqrt{(37-25)^2 + (142-40)^2}$$

$$= \sqrt{12^2 + 102^2} = \sqrt{144 + 10404} = \sqrt{10548}$$

$$= 102.70$$

$$d(a_1, b_2) = \sqrt{(37-35)^2 + (142-60)^2} = \sqrt{2^2 + 82^2} = \sqrt{6868}$$

$$\approx 82.87$$

$$d(a_1, b_3) = \sqrt{(37-45)^2 + (142-80)^2} = \sqrt{8^2 + 62^2} = \sqrt{3908} = 62.51$$

$$d(a_1, b_4) = \sqrt{(37-20)^2 + (142-20)^2} = \sqrt{17^2 + 122^2} = \sqrt{15173} \approx 123.17$$

$$d(a_1, b_5) = \sqrt{(37-35)^2 + (142-120)^2} = \sqrt{2^2 + 22^2} = \sqrt{488} \approx 22.09$$

$$d(a_1, b_6) = \sqrt{(37-52)^2 + (142-18)^2} = \sqrt{15^2 + 124^2} = \sqrt{15601}$$

$$\approx 124.90$$

$$d(a_1, b_7) = \sqrt{(37-23)^2 + (142-95)^2} = \sqrt{14^2 + 47^2} = \sqrt{2405} \approx 49.04$$

$$d(a_1, b_8) = \sqrt{(37-40)^2 + (142-62)^2} = \sqrt{3^2 + 80^2} = \sqrt{6409} \approx 80.05$$

$$d(a_1, b_9) = \sqrt{(37-60)^2 + (142-100)^2}$$

$$= \sqrt{23^2 + 42^2} = \sqrt{529 + 1764} \approx 47.88$$

$$d(a_1, b_{10}) = \sqrt{(37-48)^2 + (142-220)^2} = \sqrt{11^2 + 78^2}$$

$$= \sqrt{6205} = 78.77$$

$$d(a_1, b_{11}) = \sqrt{(37-33)^2 + (150-142)^2} = \sqrt{4^2 + 8^2} \approx 8.94$$

For k nearest neighbours algorithm, we calculate the distance of each new point from all other point & choose the k ~~closest~~ closest points.

for k = 1 : We need to choose 1 closest point to $a_1 = (37, 142)$

Req. point is → $b_{11} = (33, 150)$ with a distance of 8.94 .

So, the HPI of $(37, 142)$ = HPI of b_{11} = 264

BHK of $(37, 142)$ = BHK of b_{11} = 4

req. (HPI, BHK) = (264, 4)

for k = 2 : 2 closest points to a_1 → (33, 150) & (35, 120)

(HPI, BHK) of (33, 150) = (264, 4)

(HPI, BHK) of (35, 120) = (139, 4)

(HPI, BHK) of a_1 is predicted to be = $\left(\frac{264+139}{2}, \frac{4+4}{2}\right)$

$= \left(\frac{403}{2}, \frac{8}{2}\right) = (201.5, 4)$

The req. HPI & BHK for a_1 for k = 2 is →

HPI = 201.5 , BHK = 4

for k = 3 : 3 closest points to a_1 → (33, 150) , (35, 120) & (60, 100).

(HPI, BHK) of (33, 150) → (264, 4)

(HPI, BHK) of (35, 120) → (139, 4)

(HPI, BHK) of (60, 100) → (139, 2)

Req. HPI & BHK → $\left(\frac{264+139+139}{3}\right)$ → $\frac{542}{3} = 180.66$

(180.66, 4) ← Ans

BHK 4 appears two times where BHK 2 appears once, So BHK = 4