ML ASSIGNMENT # 01

23K-0842

Q2) $x = [x_1, x_2]^T$ $\quad l = \begin{cases} 1 & \text{if } x \text{ is positive family oriented apartment} \\ 0 & \text{if } x \text{ is negative non family oriented apartment} \end{cases}$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

A) $D = \{ (x^{(i)}, l^{(i)}) \mid i = 1, 2, 3, \ldots, M \}$

where

$$x^{(i)} = \begin{bmatrix} x_1^{(i)} \\ x_2^{(i)} \end{bmatrix}, \quad l^{(i)} \in \{0, 1\}$$

B. $H(x) = \begin{cases} 1 & , \text{if } \theta_0 + \theta_1 x_1 + \theta_2 x_2 \geq 0 \\ 0 & , \text{otherwise} \end{cases}$

where

θ_0 = bias or intercept

θ_1 = weight for x_1

θ_2 = weight for x_2

Decision boundary is the line $\theta_1 x_1 + \theta_2 x_2 + \theta_0 = 0$.

C) Hypothesis error :-

0/1 Loss function :-

$$L_{0/1} = \frac{1}{M} \sum_{i=1}^{M} (H(x_{(i)}) \neq l_{(i)})$$

$$L_{0/1}(H(x_{(i)}), l_{(i)}) = \begin{cases} 0 & \text{if } H(x_{(i)}) = l_{(i)} \\ 1 & \text{if } H(x_{(i)}) \neq l_{(i)} \end{cases}$$

Mean Squared Error :-

$$L_{MSE} = \frac{1}{M} \sum_{i=1}^{M} (H(x_{(i)}) - l_{(i)})^2$$

Root Mean Squared Error :-

$$L_{RMSE} = \sqrt{\frac{1}{M} \sum_{i=1}^{M} [H(x_{(i)}) - l_{(i)})^2}$$

CS CamScanner

D) **Most specific hypothesis :-**

It classifies only the exact positive training example as positive, (no generalization beyond them). It is too restrictive.

Most general hypothesis :-

It classifies the largest possible set as positive. It accepts everything, no constraints. It will have zero false negatives but many false positives.

Q4)

id	Price x_1 (PKR)	size x_2 (sq.ft)	Label(l)
A	55	450	Standard (0)
B	75	900	luxury (1)
C	65	500	standard (0)
D	85	850	luxury (1)
E	45	400	standard (0)

$$\text{labels} = \begin{cases} \text{standard}, & l_{(i)} = 0 \\ \text{luxury}, & l_{(i)} = 1 \end{cases}$$

$$x_{new} = (70, 700)$$

Chebyshev Distance (L_∞ Norm) :-

$$d_\infty(x,y) = \max(|x_i - y_i|) = \max(|x_1 - y_1|, |x_2 - y_2|)$$

$$d_\infty(x_{new}, A) = \max(|70-55|, |700-450|) = \max(15, 250) = 250$$

$$d_\infty(x_{new}, B) = \max(|70-75|, |700-900|) = \max(5, 200) = 200$$

$$d_\infty(x_{new}, C) = \max(|70-65|, |700-500|) = \max(5, 200) = 200$$

$$d_\infty(x_{new}, D) = \max(|70-85|, |700-850|) = \max(15, 150) = 150$$

$$d_\infty(x_{new}, E) = \max(|70-45|, |700-400|) = \max(25, 300) = 300$$

With $K = 3$, B, C, D are nearest neighbours.

luxury standard luxury $\xrightarrow{\text{Majority Voting}}$ x_{new} is Luxury(1)

City Block | Manhattan (L1 Norm):-

$$d(x,y) = \sum_{i=1}^{m} |x_i - y_i|$$

$d(x_{new}, A) = |70-55| + |700-450| = 15 + 250 = 265$

$d(x_{new}, B) = |70-75| + |700-900| = 5 + 200 = 205$

$d(x_{new}, C) = |70-65| + |700-500| = 5 + 200 = 205$

$d(x_{new}, D) = |70-85| + |700-850| = 15 + 150 = 165$

$d(x_{new}, E) = |70-45| + |700 - \overset{4}{8}00| = 25 + 300 = 325$

With $k=3$, B, C, D are nearest
luxury standard luxury → majority voting xnew is luxury (1).

Computing weights:-

$$W_i = \frac{1}{d_i}$$

Chebyshev Weights:-

$w_A = \frac{1}{250} = 0.00400 \rightarrow$ Standard (0)

$w_B = \frac{1}{200} = 0.00500 \rightarrow$ luxury (1) ✓

$w_C = \frac{1}{200} = 0.00500 \rightarrow$ Standard (0) ✓

$w_D = \frac{1}{150} \approx 0.00667 \rightarrow$ luxury (1) ✓

Nearest

$w_E = \frac{1}{300} \approx 0.00333 \rightarrow$ Standard (0)

$K=3$, so nearest were B, C, D

sum weights per class:-

- Luxury (1) : $0.00667 + 0.005 = 0.01167$
- Standard (0) : 0.005

Luxury (0.01167) > Standard (0.005), so xnew can be classified as luxury (1).

CS CamScanner

Manhattan weights :-

$w_A = \frac{1}{265} = 0.00377 \rightarrow$ Standard (0)

$w_B = \frac{1}{205} = 0.00488 \rightarrow$ Luxury (1)

$w_C = \frac{1}{205} = 0.00488 \rightarrow$ Standard (0)

$w_D = 1/165 = 0.00606 \rightarrow$ luxury (1)

$w_E = 1/325 = 0.00308 \rightarrow$ Standard (0)

D, C, B $\rightarrow$ nearest neighbours

Sum :-

- luxury (1) :- 0.00488 + 0.00606 = 0.01094
- Standard (0) :- 0.00488

luxury (1) = 0.01094 > standard (0) = 0.00488 so x_{new} is luxury (1).

Final Answers :-

- Using Manhattan (city-block), weighted KNN with k = 3 predicts ~~k = 3 predicts~~ luxury (1).
- using Chebyshev (L_∞), weighted KNN with k = 3 predicts Luxury (1).

CamScanner