• Problem-4 : K-Nearest Neighbour

(1) 5-nearest neighbour

Part-I

The given points in the graph are-

Points	x	y	class	distance from I
A	1	1	−	5
B	1	3	−	4.123
C	2	2	−	3.605
D	2	6	−	3.605
E	2.	8	+	5
F	4	1	−	3.162
G	4	4	−	1
H	5	1	−	3
I	5	4	Test	Q
J	5	9	+	5
K	6	5	+	1.414
L	6	8	+	4.123

M	7	3	+	2.236
N	7	7	+	3.605
O	8	1	—	4.242
P	8	7	+	4.242
Q	8	9	+	5.830
R	9	1	—	5
S	9	4	+	4
T	9	6	+	4.472

The formula for euclidean distance,

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Now, euclidean distance,

$A = \sqrt{(5-1)^2 + (4-1)^2} = 5$

$B = \sqrt{(5-1)^2 + (4-3)^2} = 4.123$

$C = \sqrt{(5-2)^2 + (4-2)^2} = 3.605$

$D = \sqrt{(5-2)^2 + (4-6)^2} = 3.605$

$E = \sqrt{(5-2)^2 + (4-8)^2} = 5$

$F = \sqrt{(5-4)^2 + (4-1)^2} = 3.162$

$G = \sqrt{(5-4)^2 + (4-4)^2} = 1$

$H = \sqrt{(5-5)^2 + (4-1)^2} = 3$

$J = \sqrt{(5-5)^2 + (4-9)^2} = 5$

$K = \sqrt{(5-6)^2 + (4-5)^2} = 1.414$

$L = \sqrt{(5-6)^2+(4-8)^2} = 4.123$

$M = \sqrt{(5-7)^2+(4-3)^2} = 2.236$

$N = \sqrt{(5-7)^2+(4-7)^2} = 3.605$

$O = \sqrt{(5-8)^2+(4-1)^2} = 4.242$

$P = \sqrt{(5-8)^2+(4-7)^2} = 4.242$

$Q = \sqrt{(5-8)^2+(4-9)^2} = 5.830$

$R = \sqrt{(5-9)^2+(4-1)^2} = 5$

$S = \sqrt{(5-9)^2+(4-4)^2} = 4$

$T = \sqrt{(5-9)^2+(4-6)^2} = 4.472$

So, from the distance, the following 5 are the nearest to point I,

Point	(x, y)	class
F	(4, 1)	−
G	(4, 4)	−
H	(5, 1)	−
K	(6, 5)	+
M	(7, 3)	+

Since, the majority of the point class is '−', the test point I will be classified '−'.

(2) Manhattan distance weighted 3-nearest neighbors.
(The weight is $1/d^2$)

⇒ The Manhattan distance is given by

$$d(x_i, x_j) = \sum_{m=1}^{D} |x_{im} - x_{jm}|$$

Using the points table from the previous sub-question,

	weight = $1/d^2$
A = $\|5-1\| + \|4-1\| = 4+3 = 7$	0.0204
B = $\|5-1\| + \|4-3\| = 4+1 = 5$	0.04
C = $\|5-2\| + \|4-2\| = 3+2 = 5$	0.04
D = $\|5-2\| + \|4-6\| = 3+2 = 5$	0.04
E = $\|5-2\| + \|4-8\| = 3+4 = 7$	0.0204
F = $\|5-4\| + \|4-1\| = 1+3 = 4$	0.0625
G = $\|5-4\| + \|4-4\| = 1+0 = 1$	1
H = $\|5-5\| + \|4-1\| = 0+3 = 3$	0.111
I = $\|5-5\| + \|4-9\| = 0+5 = 5$	0.04
J = $\|5-6\| + \|4-5\| = 1+1 = 2$	0.25
K = $\|5-6\| + \|4-8\| = 1+4 = 5$	0.04
L = $\|5-7\| + \|4-3\| = 2+1 = 3$	0.111
M = $\|5-7\| + \|4-7\| = 2+3 = 5$	0.04
N = $\|5-8\| + \|4-7\| = 3+3 = 6$	0.0277
O = $\|5-8\| + \|4-7\| = 3+3 = 6$	0.0277
P = $\|5-8\| + \|4-9\| = 3+5 = 8$	0.015
Q = $\|5-9\| + \|4-1\| = 4+3 = 7$	0.0204
R = $\|5-9\| + \|4-4\| = 4+0 = 4$	0.0625
S = $\|5-9\| + \|4-6\| = 4+2 = 6$	0.0277

⇒ So, from manhattan distance, we get the 3 - nearest neighbors are -

Point	(x, y)	class	weight
G	(4, 4)	−	1
H	(5, 1)	−	0.111
K	(6, 5)	+	0.25
M	(7, 3)	+	0.1

⇒ As, it is equally distributed between -ve and +ve class, we will check the weights

Therefore, point G has the highest weight, hence, the test point I will be classified as '−' Negative.