Anomaly Detection Activity

x	y
1	1
1	1.5
2	1
5	4
0.5	1
0.5	0.5

For simplicity, first, we can calculate the Euclidean distance of each points to others:

	P1	P2	P3	P4	P5	P6
P1=(1,1)	0	0.5	1.0	5.0	0.5	0.7
P2=(1,1.5)	0.5	0	1.1	4.7	0.7	1.1
P3=(2,1)	1.0	1.1	0	4.2	1.5	1.6
P4=(5,4)	5.0	4.7	4.2	0	5.4	5.7
P5=(0.5,1)	0.5	0.7	1.5	5.4	0	0.5
P6=(0.5,0.5)	0.7	1.1	1.6	5.7	0.5	0

Let's calculate the anomaly score for P1=(1,1) when K=2. For the other points the procedure is the same:

K = 2

A)

d = 0.5 d=0.5

1) Calculate 2-NN list of P1 = [P2=(1,1.5), P5=(0.5,1)]

2) Calculate density of (1,1) based definition (A):
   Density of P1 = 1/0.5 = 2.0 (0.5 is distance to P5)

3) Calculate the density of elements in the list [P2=(1,1.5), P5=(0.5,1)] based on definition (A):

   d = 0.5 d=0.7
   2-NN list of P2 = [P1=(1,1), P5=(0.5,1)]
   Density of P2 = 1/0.7 = 1.4 (0.7 is distance to P5)

   d = 0.5 d=0.5
   2-NN list of P5 = [P1=(1,1), P6=(0.5,0.5)]
   Density of P5 = 1/0.5 = 2.0 (0.5 is distance to P6)

   Relative density = ((1.4+2.0)/2)/2.0 = 0.85

K = 2

B)

Let's calculate the anomaly score for P1=(1,1). For other points the procedure is the same:

d = 0.5 d=0.5

1) Calculate 2-NN list of P1 = [P2=(1,1.5), P5=(0.5,1)]

2) Calculate density of P1 based definition (B):
   Density of (1,1) = 1/Average(0.5,0.5) = 2.0

3) Calculate the density of elements in the list [P2=(1,1.5), P5=(0.5,1)] based on definition (B):

   d = 0.5 d=0.7
   2-NN list of P2 = [P1=(1,1), P5=(0.5,1)]
   Density of P2 = 1/Average(0.5,0.7) = 1.7

   d = 0.5 d=0.5
   2-NN list of P5 = [P1=(1,1), P6=(0.5,0.5)]
   Density of P5 = 1/Average(0.5,0.5) = 2.0

   Relative density = ((1.7+2.0)/2)/2.0 = 0.93