Cluster points

Student:10428143 簡琪臻 Date：2018/06/10 Adviser: 林康平

School Department: 中原大學電機工程學系

Abstract

Use Forgy's Algorithm to cluster points with k=2, k=4, and k=6 separately, and analyze the results.

Introduction

There are many different algorithms to cluster things, the most common methods are hierarchical clustering and partitional clustering. This paper use Forgy's Algorithm in partitional clustering to cluster points.

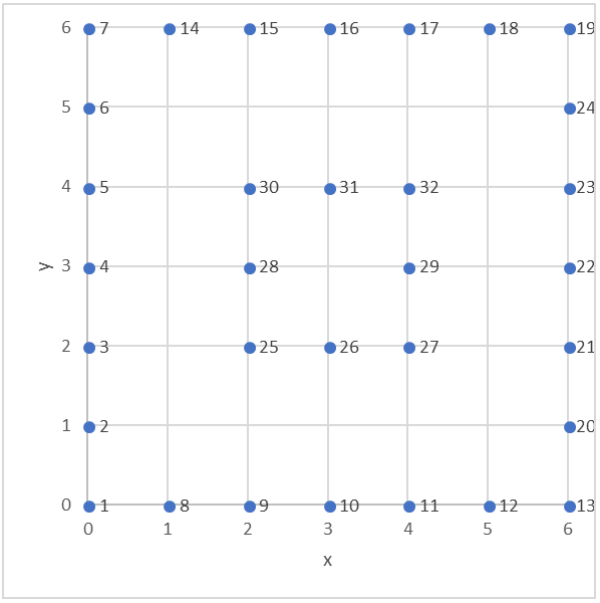
Method

1. Give every point a coordinate and a number.
2. Use Forgy's Algorithm to cluster points with k=2, k=4, and k=6 separately and record the results.

Result

1. Data

number	x	y	number	x	y
1	0	0	17	4	6
2	0	1	18	5	6
3	0	2	19	6	6
4	0	3	20	6	1
5	0	4	21	6	2
6	0	5	22	6	3
7	0	6	23	6	4
8	1	0	24	6	5
9	2	0	25	2	2
10	3	0	26	3	2
11	4	0	27	4	2
12	5	0	28	2	3
13	6	0	29	4	3
14	1	6	30	2	4
15	2	6	31	3	4
16	3	6	32	4	4



2. Forgy's Algorithm
   a. Set k=2, which will produce two clusters, and use the 1(0, 0) and 24 (6, 5) in the list as seed points.
      i. The clusters {1,2,3,4,5,6,7,8,9,10,11,12,25,26,28}, and {13,14,15,16,17,18,19,20,21,22,23,24,27,29,30,31,32} are produced.

Sample	Nearest cluster centroid	Sample	Nearest cluster centroid
1	(0, 0)	17	(6, 5)
2	(0, 0)	18	(6, 5)
3	(0, 0)	19	(6, 5)
4	(0, 0)	20	(6, 5)
5	(0, 0)	21	(6, 5)
6	(0, 0)	22	(6, 5)
7	(0, 0)	23	(6, 5)
8	(0, 0)	24	(6, 5)
9	(0, 0)	25	(0, 0)
10	(0, 0)	26	(0, 0)

11	(0, 0)	27	(6, 5)
12	(0, 0)	28	(0, 0)
13	(6, 5)	29	(6, 5)
14	(6, 5)	30	(6, 5)
15	(6, 5)	31	(6, 5)
16	(6, 5)	32	(6, 5)

ii. The centroid of the clusters changes to (1.5, 1.9) and (4.4, 4), and the clusters {1,2,3,4,5,6,7,8,9,10,11,12,25,26,28,30} , and {13,14,15,16,17,18,19,20,21,22,23,24,27,29,31,32} are produced.

Sample	Nearest cluster centroid	Sample	Nearest cluster centroid
1	(1.5, 1.9)	17	(4.4, 4)
2	(1.5, 1.9)	18	(4.4, 4)
3	(1.5, 1.9)	19	(4.4, 4)
4	(1.5, 1.9)	20	(4.4, 4)
5	(1.5, 1.9)	21	(4.4, 4)
6	(1.5, 1.9)	22	(4.4, 4)
7	(1.5, 1.9)	23	(4.4, 4)
8	(1.5, 1.9)	24	(4.4, 4)
9	(1.5, 1.9)	25	(1.5, 1.9)
10	(1.5, 1.9)	26	(1.5, 1.9)
11	(1.5, 1.9)	27	(4.4, 4)
12	(1.5, 1.9)	28	(1.5, 1.9)
13	(4.4, 4)	29	(4.4, 4)
14	(4.4, 4)	30	(1.5, 1.9)
15	(4.4, 4)	31	(4.4, 4)
16	(4.4, 4)	32	(4.4, 4)

iii. The centroid of the clusters changes to (1.5, 2) and (4, 4), and the clusters {1,2,3,4,5,6,7,8,9,10,11,12,25,26,28,30} , and {13,14,15,16,17,18,19,20,21,22,23,24,27,29,31,32} are produced, no clusters changed, so the algorithm terminates.

b. Set k=4, which will produce four clusters, and use the 8(1, 0), 16(3, 6), 24(6, 5), and 32(4, 4) in the list as seed points.

i. The clusters {1,2,3,4,8,9,10,11,12,25}, and {5,6,7,14,15,16,17}, {18,19,22,23,24}, and {13,20,21,26,27,28,29,30,31,32} are produced.

Sample	Nearest cluster centroid	Sample	Nearest cluster centroid
1	(1, 0)	17	(3, 6)
2	(1, 0)	18	(6, 5)
3	(1, 0)	19	(6, 5)
4	(1, 0)	20	(4, 4)
5	(3, 6)	21	(4, 4)
6	(3, 6)	22	(6, 5)
7	(3, 6)	23	(6, 5)
8	(1, 0)	24	(6, 5)
9	(1, 0)	25	(1, 0)
10	(1, 0)	26	(4, 4)
11	(1, 0)	27	(4, 4)
12	(1, 0)	28	(4, 4)
13	(4, 4)	29	(4, 4)
14	(3, 6)	30	(4, 4)
15	(3, 6)	31	(4, 4)
16	(3, 6)	32	(4, 4)

ii. The centroid of the clusters changes to (1.7, 0.8), (1.4, 5.6), (5.8, 4.8) and (4, 2.5), and the clusters {1,2,3,4,8,9,10,11,25}, and {5,6,7,14,15,16,30}, {17,18,19,22,23,24}, and {12,13,20,21,26,27,28,29,31,32} are produced.

iii. The centroid of the clusters changes to (1.3, 0.9), (1.1, 5.3), (5.5, 5) and (4.3, 2.1), and the clusters {1,2,3,4,8,9,10,25,28}, and {5,6,7,14,15,16,30,31}, {17,18,19,23,24,32}, and {11,12,13,20,21,22,26,27,29} are

produced.

iv. The centroid of the clusters changes to (1.11, 1.22), (1.375, 5.125), (5.167, 5.167) and (4.89, 1.44), and the clusters {1,2,3,4,8,9,10,25,28}, {5,6,7,14,15,16,30,31}, {17,18,19,23,24,32}, and {11,12,13,20,21,22,26,27,29} are produced, no clusters changed, so the algorithm terminates.

c. Set k=6, which will produce six clusters, and use the 5(0, 4), 10(3, 0), 15(2, 6), 20(6, 1), 25(2, 2), 30(2, 4) and in the list as seed points.

i. The clusters {3,4,5,6,7}, {8,9,10,11}, {14,15,16,17,18,19}, {12,13,20,21,22,23,24}, {1,2,25,26,27,28,29}, and {30,31,32} are produced.

Sample	Nearest cluster centroid	Sample	Nearest cluster centroid
1	(2, 2)	17	(2, 6)
2	(2, 2)	18	(2, 6)
3	(0, 4)	19	(2, 6)
4	(0, 4)	20	(6, 1)
5	(0, 4)	21	(6, 1)
6	(0, 4)	22	(6, 1)
7	(0, 4)	23	(6, 1)
8	(3, 0)	24	(6, 5)
9	(3, 0)	25	(2, 2)
10	(3, 0)	26	(2, 2)
11	(3, 0)	27	(2, 2)
12	(6, 1)	28	(2, 2)
13	(6, 1)	29	(2, 2)
14	(2, 6)	30	(2, 4)
15	(2, 6)	31	(2, 4)
16	(2, 6)	32	(2, 4)

ii. The centroid of the clusters changes to (0, 4), (2.5, 0), (3.5, 6), (5.9, 2.1), (2.1, 1.9), (3, 4) and the clusters {3,4,5,6,7,14}, {1,8,9,10,11}, {15,16,17,18,19,24}, {12,13,20,21,22,23,27}, {2,25,26,28}, and{29,30,31,32} are produced.

iii. The centroid of the clusters changes to (0.17, 4.33), (2, 0), (4.33, 5.83), (5.57, 1.71), (2.2, 2), (3.25, 3.75) and the clusters {4,5,6,7,14}, {1,2,8,9,10,11}, {15,16,17,18,19,24}, {12,13,20,21,22,23,27}, {3,25,26,28}, and{29,30,31,32} are produced.

iv. The centroid of the first, second, and fifth clusters changes to (0.2, 4.8), (1.67, 0.17), (2.2, 2.2) and the clusters {4,5,6,7,14,15}, {1,2,8,9,10}, {16,17,18,19,24}, {11,12,13,20,21,22,23,27}, {3,25,26,28}, and{29,30,31,32} are produced.

v. The centroid of the first, second, third, and fourth clusters changes to (0.5, 5), (1.2, 0.2), (4.8, 5.8), (5.375, 1.5) and the clusters {4,5,6,7,14,15}, {1,2,3,8,9,10}, {16,17,18,19,23,24}, {11,12,13,20,21,22,27}, {25,26,28}, and{29,30,31,32} are produced.

vi. The centroid of the second, third, fourth, and fifth clusters changes to (1, 0.5), (5, 5.5), (5.28, 1.14), (2.33, 2.33) and the clusters {4,5,6,7,14,15}, {1,2,3,8,9,10}, {16,17,18,19,23,24}, {11,12,13,20,21,22,27}, {25,26,28}, and{29,30,31,32} are produced, no clusters changed, so the algorithm terminates.

Discussion and Conclusion

1. The values of seed points which we chose at the first are important that they may influence the time of iterations the algorithm took.
2. Even the values of seed points are biased, the content of each cluster can reach even eventually.
3. Basically, the time of iterations increases with the value of k.

References

Textbook