July
WEEK 28

May 2012 | June 2012 | July 2012

12 Thursday 194/172

Orangeman's Day (N Ireland-U.K.)

8.00 am

suppose the data mining task is to cluster point (with x, y representing location) into three cluster where point are.

A1 (2,10) B1 (5, 8) a(1, 2)

A2 (2,5) B2(7,5) c2(4,9)

A3 (8,4) B3 (6,4)

The distance funtion is Euclidean distance. supose initialy we assign A1, B1, e1 as center of each cluster respectively use the kmean.

algorith to show only

a) The three cluster centers after the first round of execution

b) the final three cluster.

pm

July
WEEK 27

May 2012

M		7	14	21	28
T	1	8	15	22	29
W	2	9	16	23	30
T	3	10	17	24	31
F	4	11	18	25	
S	5	12	19	26	
S	6	13	20	27	

June 2012

M		4	11	18	25
T		5	12	19	26
W		6	13	20	27
T		7	14	21	28
F	1	8	15	22	29
S	2	9	16	23	30
S	3	10	17	24	

July 2012

M	30	2	9	16	23
T	31	3	10	17	24
W		4	11	18	25
T		5	12	19	26
F		6	13	20	27
S		7	14	21	28
S	1	8	15	22	29

6 Friday 188/178

8.00 am The distance function between two
8.30 point $a=(x_1,y_1)$ and $b=(x_2,y_2)$ is
9.00 defined as.

$$d(p,q) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}$$

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

Iteration ①

We calculate the distance of each point from each of the center of the three clusters. by using Euclidean funtion.

$A_1(2,10)$ $A_2(2,5)$, $A_3(8,4)$, $B_1(5,8)$, $B_2(7,5)$
$B_3(6,4)$, $C_1(1,2)$ $C_2($

	A1	A2	A3	B1	B2	B3	C1	C2
C1:A1	0	50	8.48	3.60	7.07	7.21	8.06	2.23
C2: B1	3.6	4.24	5.0	0	3.6	4.12	7.21	1.41
C3: C1	8.06	3.16	7.28	7.21	6.7	5.38	0	7.61

the three clusters with cluster points are.

cluster 1 = {A1(2,10)}

cluster 2 = {A3(8,4), B1(5,8), B2(7,5), B3(6,4), C2(4,9)}

cluster 3 = {A2(2,5), C1(4,9)}

lets calulater centr after fist round.

center 1 = (2,10)

center 2 = {(5+8+7+6,4)/5, (8+4+5+4+9)/

= (6,6)

center 3 = (1.5, 3.5)

187/179 Thursday 5

The final three cluster
Second iteration.

	A1	A2	A3	B1	B2	B3	C1	C2
centroid 1: A1	0	5	8.48	3.6	7.07	7.21	8.06	2.23
centroid 2: B1	4.12	4.12	2.82	2.23	1.41	2	6.4	3.6
centroid 3: C1	6.51	1.58	6.51	5.7	5.7	4.52	1.58	0

After third i

After third iteration the final cluster are

cluster 1: {(A1, C2, B1)}

cluster 2: {(A3, B2, B3)}

cluster 3: {(A2, C1)}