

مسئله ۱) از فاصله اقلیدسی برای محاسبه فاصله ها استفاده می کنیم:  
 $d = \sqrt{(y_2 - y_1)^2 + (x_2 - x_1)^2}$  (برای دو نقطه)

الف)

نقاط	مرکز cluster ها			شماره cluster ها
	$C_1(2,10)$	$C_2(5,8)$	$C_3(1,2)$	
$A_1(2,10)$	0	3.61	8.06	1
$A_2(2,6)$	5	4.24	3.16	3
$A_3(8,4)$	8.49	5	7.28	2
$A_4(5,8)$	3.61	0	7.21	2
$A_5(7,5)$	7.07	3.61	6.71	2
$A_6(6,4)$	7.21	4.12	5.39	2
$A_7(1,2)$	8.06	7.21	0	3
$A_8(4,9)$	2.24	1.41	7.62	2

مرحله ۱)

کاری که باید انجام دهیم، فاصله اقلیدسی هر نقطه تا هر مرکز را بدست آوردن و در خانه متناظر آن می نویسیم.

سپس برای هر نقطه (هر سطر)، cluster ای که آن نقطه درون آن قرار می گیرد همان cluster ای

است که کمترین فاصله را تا مرکز آن دارد. سپس از تکرار کردن این مقایسه در جدول، برای تعیین مرکز clusterهای

جدید برای مرحله بعدی، میان نقاط موجود در cluster مربوط به خوشه های کاملاً ثابت گرفته و نقطه بدست آمده،

می شود مرکز جدید مربوط به آن شماره cluster محاسبات را با استفاده از ماشین حساب انجام داد و

از نوشتن آنها در اینجا خودداری کردم به دلیل طولانی بودن

Subject: \_\_\_\_\_

Year: \_\_\_\_\_

Month: \_\_\_\_\_

Date: \_\_\_\_\_

( )

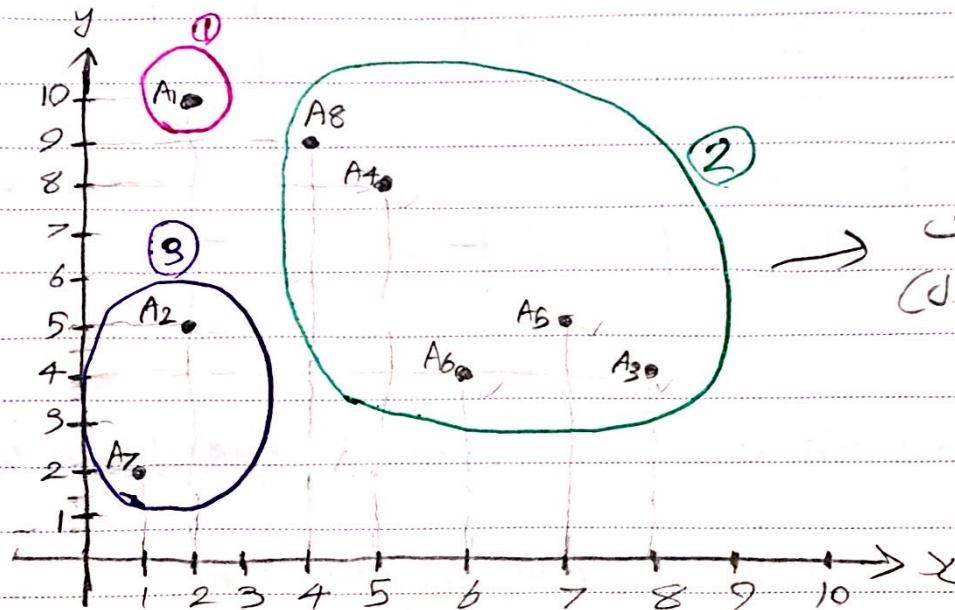
ب) زیرا فقط یک نقطه درونش بود  
→  
cluster 1: (2, 10)

بهترین مرکز جدید:

$$\text{cluster 2: } x = \frac{8+5+7+6+4}{5} = 6, \quad y = \frac{4+8+5+4+9}{5} = 6 \rightarrow (6, 6)$$

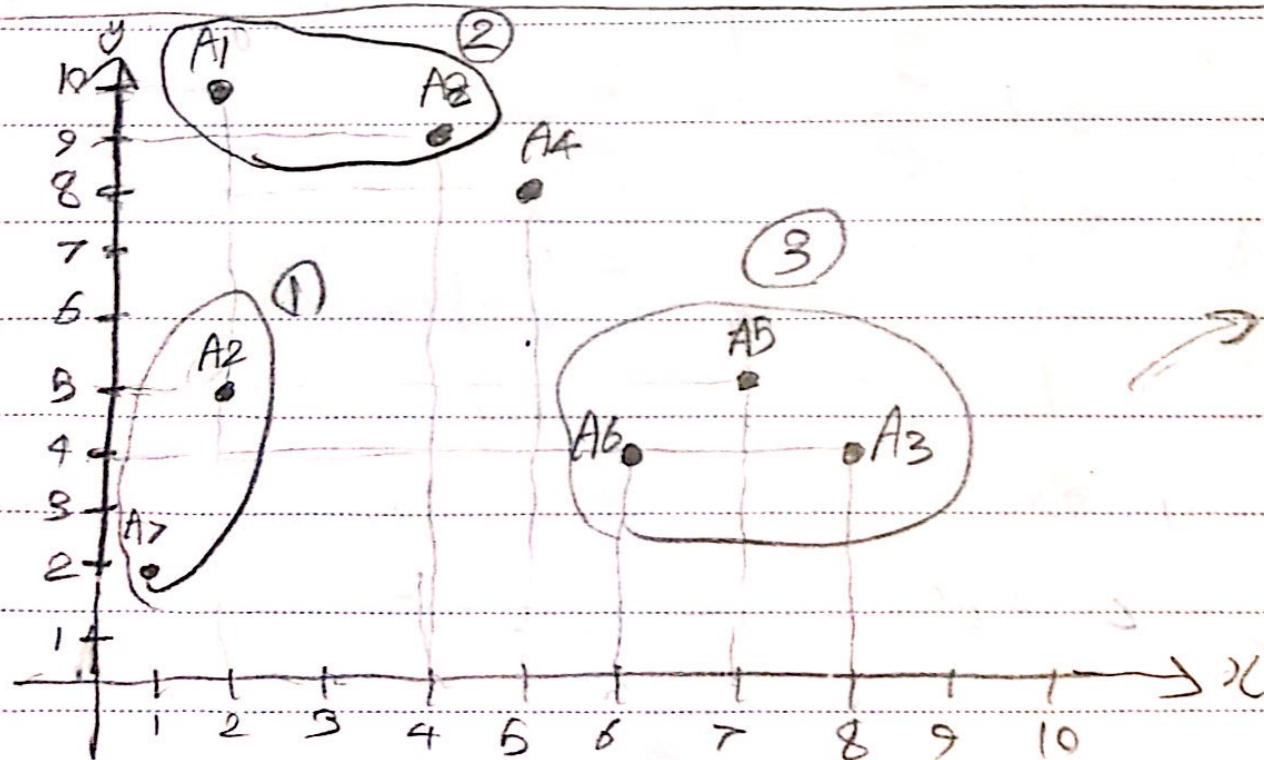
$$\text{cluster 3: } x = \frac{2+1}{2} = 1.5, \quad y = \frac{5+2}{2} = 3.5 \rightarrow (1.5, 3.5)$$

ج)



→ کلاسهای برای قسمت  
الف (epoch اول)





مجموعه های داده



Sunwood

Q

پیرس محاسبہ لین و فیکس ، اس قدر لین حلقہ را تکراری کنیم و جدول می کشیم تا نقاط دو مرحله سوال تغییر cluster

نکته: پس شروع می کنیم با مرکز cluster های جدید که از قسمت قبل بدست آوریم ، مرحله پنجم را انجام می دهیم:

Sunwood

Subject: \_\_\_\_\_

Year: \_\_\_\_\_ Month: \_\_\_\_\_ Date: \_\_\_\_\_ ( )

مرحله 2

نقاط	مرکز cluster ها			شماره cluster
	$C_1(2,10)$	$C_2(6,6)$	$C_3(1.5,3.5)$	
$A_1$	0	5.66	6.52	1
$A_2$	5	4.12	1.58	3
$A_3$	8.99	2.83	6.52	2
$A_4$	3.61	2.24	5.7	2
$A_5$	7.07	1.41	5.7	2
$A_6$	7.21	2	4.53	2
$A_7$	8.06	6.4	1.58	3
$A_8$	2.24	3.61	6.04	1

شماره cluster برای  $A_8$  →  
 نسبت به مرحله قبلی تغییر کرده و پس باید داده به دسته 1  
 برده ایم. (مرکز کلاسترهای جدید را مانند بخش قبلی محاسبه می کنیم)

مرحله 3

نقاط	مرکز cluster ها			شماره cluster
	$C_1(3,9.5)$	$C_2(6.5,5.25)$	$C_3(1.5,3.5)$	
$A_1$	1.12	6.54	6.52	1
$A_2$	4.61	4.51	1.58	3
$A_3$	7.43	1.95	6.52	2
$A_4$	2.5	3.13	5.7	1
$A_5$	6.02	0.56	5.7	2
$A_6$	6.26	1.35	4.53	2
$A_7$	7.76	6.39	1.58	3
$A_8$	1.12	4.51	6.04	1

شماره cluster برای  $A_4$  →  
 تغییر کرده و پس باید داده به دسته 1



Subject: \_\_\_\_\_

Year. \_\_\_\_\_ Month. \_\_\_\_\_ Date. \_\_\_\_\_ ( )

مرحله 4

نقاط	مرکز cluster ها			شماره cluster
	$C_1(3.679)$	$C_2(7.433)$	$C_3(1.5, 3.5)$	
$A_1$	1.94	7.56	6.52	1
$A_2$	4.33	5.04	1.58	3
$A_3$	6.62	1.05	6.52	2
$A_4$	1.67	4.18	5.7	1
$A_5$	5.21	0.67	5.7	2
$A_6$	5.52	1.05	4.53	2
$A_7$	7.49	6.44	1.58	3
$A_8$	0.33	5.55	6.04	1

همانطور که مشاهده می‌کنیم، شماره cluster تمام نقاط با مرحله قبلی یکسان است و در نتیجه

الگوریتم همگرا شده است. پس 3 مرحله دیگر (فاز از مرحله 1) لازم بود تا انجام شود تا حقیر بتوانم

الگوریتم همگرا شده است.

## سوال 2

ابتدا باید فاصله اولی نقاط بین هر دو نقطه را بدست آوریم و در جدول بنویسیم. سپس برای تعیین core points

باید ببینیم که برای هر نقطه، فاصله چند نقطه تا آن کمتر از  $\epsilon$  است. اگر تعداد این نقاط بزرگتر مساوی

minpts شود، آن نقطه یک core point است.

نقطه	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$	$A_7$	$A_8$
$A_1$	0							
$A_2$	5	0						
$A_3$	8.48	6.08	0					
$A_4$	3.6	4.24	5	0				
$A_5$	7.07	5	1.41	3.6	0			
$A_6$	7.21	4.12	2	4.12	1.41	0		
$A_7$	8.06	3.26	7.28	7.21	6.7	5.38	0	
$A_8$	2.23	4.47	6.4	1.41	5.38	5.38	7.61	0

فاصله هر دو نقطه را  
در جدول بنویسیم  
و بعد از آن  
ببینیم که برای  
هر نقطه، فاصله  
چند نقطه تا آن  
کمتر از  $\epsilon$  است.

حال ببینیم برای  $\epsilon = 2$  کدام نقاط core point هستند. هر نقطه که در فاصله  $\epsilon$  از آن هست

$A_1: -$  /  $A_2: -$  /  $A_3: A_5, A_6$  /  $A_4: A_8$  /  $A_5: A_3, A_6$  /  $A_6: A_3, A_5$  /

$A_7: -$  /  $A_8: A_4$

چون  $minpts = 2$  است، پس هر نقطه که حداقل با 2 نقطه دیگر (core point) در فاصله  $\epsilon$  باشد

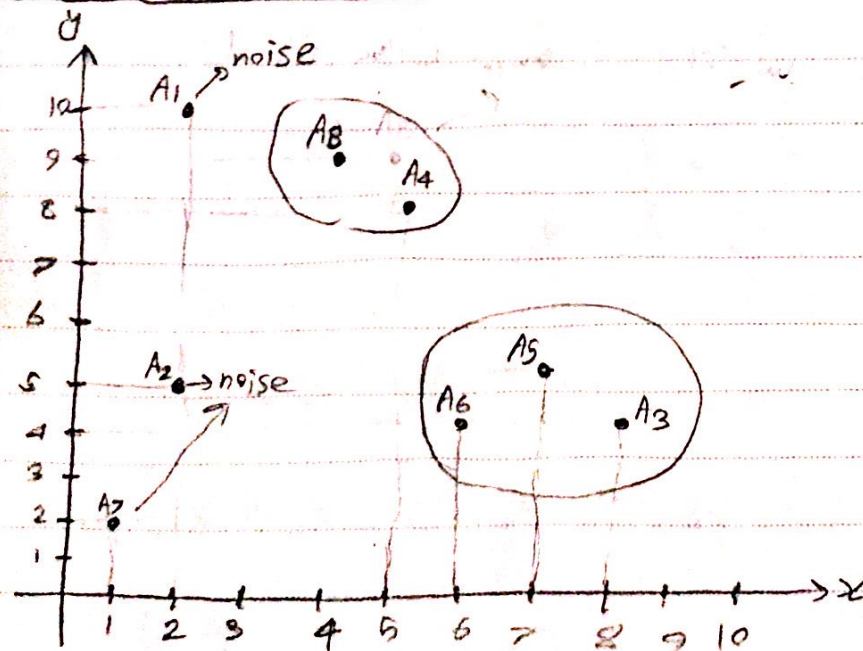
$A_3, A_4, A_5, A_6, A_8$



حال باید نحوه گردام/انتخاب را مشخص کنیم!

① core part : توسیع داده شده ② border : دایره مسایک بر یک core part ③ noise : دایره بیرون

نقطه	نوع نقطه
A <sub>1</sub>	noise
A <sub>2</sub>	noise
A <sub>3</sub>	core
A <sub>4</sub>	core
A <sub>5</sub>	core
A <sub>6</sub>	core
A <sub>7</sub>	noise
A <sub>8</sub>	core





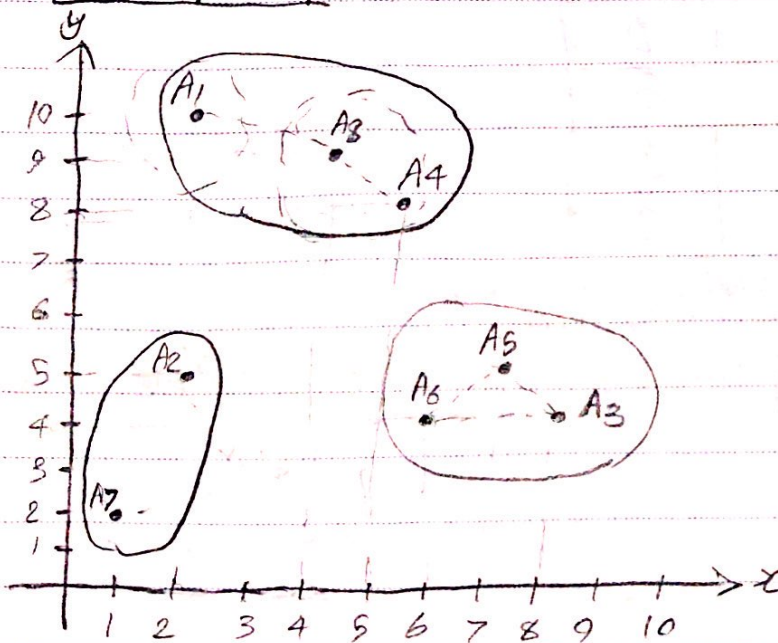
Subject: \_\_\_\_\_

Year: \_\_\_\_\_ Month: \_\_\_\_\_ Date: \_\_\_\_\_ ( )

$\rightarrow = 3.16$   
 $\therefore \epsilon = \sqrt{10}$  برای

$A_1: A_8 / A_2: A_7 / A_3: A_5, A_6 / A_4: - / A_5: A_3, A_6 / A_6: A_3, A_5 / A_7: A_2 / A_8: A_1, A_4$

نقطه	نوع نقطه
A <sub>1</sub>	Core
A <sub>2</sub>	core
A <sub>3</sub>	core
A <sub>4</sub>	border
A <sub>5</sub>	core
A <sub>6</sub>	core
A <sub>7</sub>	core
A <sub>8</sub>	core



Subject:

Year. Month. Date. ( )

سوال 3

	A	B	C	D
A	0	1	4	5
B	1	0	3	6
C	4	3	0	2
D	5	6	2	0

در هر مرحله باید کوچکترین درایه ماتریس را پیدا کرده (که در واقع کمترین فاصله بین آن دو cluster است) بگوئیم

آن دو را با هم merge کنیم و یک بزرگتر محول و عرض ماتریس کم می شود. این کار را تا زمانی که ماتریس  $1 \times 1$  شود:

مرحله 1: کمترین درایه در ماتریس بالا برای A و B است. پس با هم merge می کنیم:

	(A, B)	C	D
(A, B)	0	3	5
C	3	0	2
D	5	2	0

زیرا گفته اند که فاصله دو دسته برابر فاصله نزدیکترین نقاط آن دو دسته باشد. پس عدد 3 (فاصله B تا C) و 4 (فاصله A تا C) را انتخاب می کنیم و در این ماتریس می گذاریم.

مرحله 2: کمترین درایه مربوط به C و D است:

	(A, B)	(C, D)
(A, B)	0	3
(C, D)	3	0

مرحله 3: ترکیب (A, B) و (C, D):

	A, B, C, D
A, B, C, D	0



Subject: \_\_\_\_\_

Year. \_\_\_\_\_ Month. \_\_\_\_\_ Date. \_\_\_\_\_ ( )

سلسلہ مندرجہ ذیل hierarchical آن پر مبنی ہے:

