

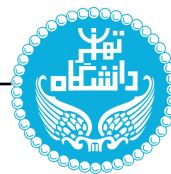
## تمرین شماره ۳

طراحان: نیما مدیر، محمدرضا علوی، محمدامین یوسفی

مهلت تحویل: شنبه ۱ اردیبهشت ۱۴۰۳، ساعت ۲۳:۵۹

## هوش مصنوعی

مدرسین: دکتر فدایی و دکتر یعقوبزاده



## بخش کتبی

### Clustering

#### سوال اول

با استفاده از الگوریتم k-means نقاط زیر را در سه cluster قرار دهید و به سوالات پاسخ دهید (فرض کنید مرکز اولیه ی cluster ها به ترتیب نقاط A1 و A4 و A7 هستند. الگوریتم را فقط یک epoch اجرا کنید):

$A1=(2,10)$ ,  $A2=(2,5)$ ,  $A3=(8,4)$ ,  $A4=(5,8)$ ,  $A5=(7,5)$ ,  $A6=(6,4)$ ,  $A7=(1,2)$ ,  $A8=(4,9)$

الف) بعد از یک epoch برای هر نقطه cluster آن را مشخص کنید.

ب) مرکز های جدید را محاسبه کنید.

ج) نقاط را رسم کنید و در شکل cluster های آن ها را مشخص کنید.

د) چند iteration دیگر نیاز است تا الگوریتم همگرا شود؟ (دیگر cluster های نقاط تغییر نکند) برای هر iteration صرفا نقاط و cluster های آن ها را رسم کنید.

### پاسخ

الف)

$A1: d(A1,C1) = 0, d(A1,C2) = \sqrt{13}, d(A1,C3) = \sqrt{65} \rightarrow A1 \text{ in cluster } 1$

$A2: d(A2,C1) = 5, d(A2,C2) = \sqrt{18}, d(A2,C3) = \sqrt{10} \rightarrow A2 \text{ in cluster } 3$

$A3: d(A3,C1) = 6, d(A3,C2) = 5, d(A3,C3) = \sqrt{53} \rightarrow A3 \text{ in cluster } 2$

$A4: d(A4,C1) = \sqrt{13}, d(A4,C2) = 0, d(A4,C3) = \sqrt{52} \rightarrow A4 \text{ in cluster } 2$

A5:  $d(A5, C1) = \sqrt{50}$ ,  $d(A5, C2) = \sqrt{13}$ ,  $d(A5, C3) = \sqrt{45} \rightarrow A5$  in cluster 2

A6:  $d(A6, C1) = \sqrt{52}$ ,  $d(A6, C2) = \sqrt{17}$ ,  $d(A6, C3) = \sqrt{29} \rightarrow A6$  in cluster 2

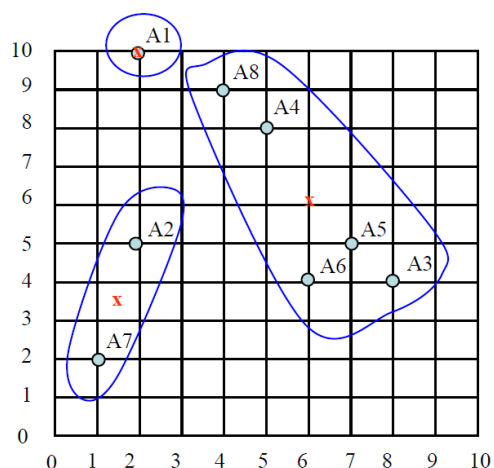
A7:  $d(A7, C1) = \sqrt{65}$ ,  $d(A7, C2) = \sqrt{52}$ ,  $d(A7, C3) = 0 \rightarrow A7$  in cluster 3

A8:  $d(A8, C1) = \sqrt{5}$ ,  $d(A8, C2) = \sqrt{2}$ ,  $d(A8, C3) = \sqrt{58} \rightarrow A8$  in cluster 2

(ب)

$C1 = (2, 10)$     $C2 = (6, 6)$     $C3 = (1.5, 3.5)$

(ج)

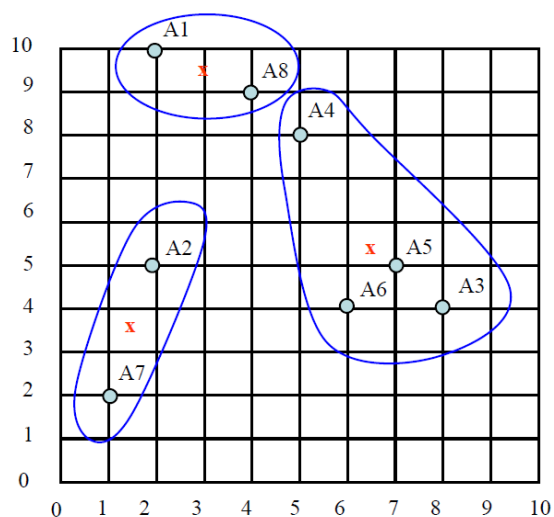


(د) دو epoch دیگر لازم است.

بعد از epoch دوم:

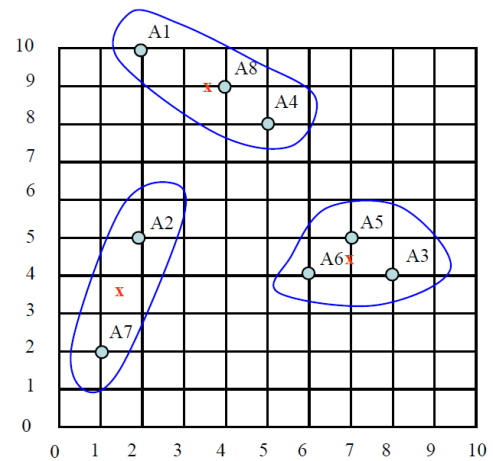
A1, A8 in cluster 1   A3, A4, A5, A6 in cluster 2   A2, A7 in cluster 3

$C1 = (3, 9.5)$     $C2 = (6.5, 5.25)$     $C3 = (1.5, 3.5)$



بعد از epoch سوم:

A1, A4, A8 in cluster 1    A3, A5, A6 in cluster 2    A2, A7 in cluster 3  
 $C1 = (3.66, 9)$      $C2 = (7, 4.33)$      $C3 = (1.5, 3.5)$



## سوال دوم

نقاط سوال قبل را با الگوریتم DBSCAN دسته بندی کنید. یک بار  $\epsilon$  را 2 و یک بار  $\sqrt{10}$  در نظر بگیرید. در هر دو حالت minpoints را 2 در نظر بگیرید. شکل نقاط و cluster ها را رسم کنید. نقاط همسایه ی هر نقطه و نقاط نويز را نیز مشخص کنید.

## پاسخ

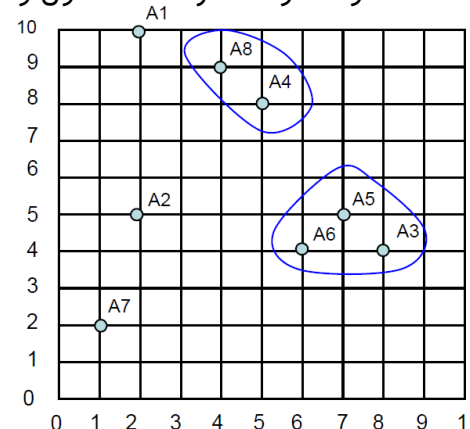
$\epsilon$  برابر 2 باشد:

نقاطی که در همسایگی هر نقطه هستند به شکل زیر است:

- A1: -
- A2: -
- A3: A5, A6
- A4: A8
- A5: A3, A6
- A6: A3, A5
- A7: -
- A8: A4

نقاط A1 و A2 و A7 نويز هستند.

نقاط A3 و A5 و A6 در cluster اول و نقاط A4 و A8 در cluster دوم هستند.



$\epsilon$  برابر  $\sqrt{10}$  باشد:

نقاطی که در همسایگی هر نقطه هست به شکل زیر است:

A1: A8

A2: A7

A3: A5, A6

A4: A8

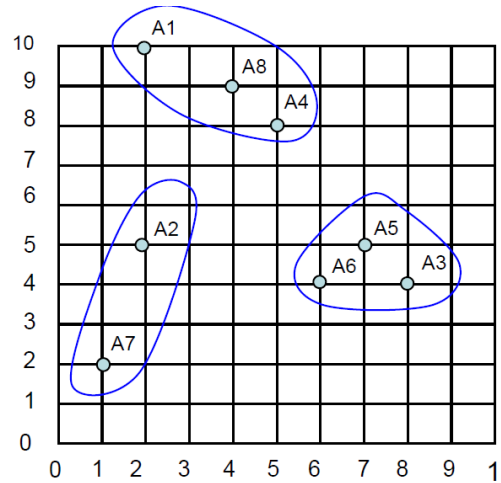
A5: A3, A6

A6: A3, A5

A7: A2

A8: A1, A4

نقاط A1 و A4 و A8 در cluster اول و نقاط A2 و A7 در cluster دوم و نقاط A3 و A5 و A6 در cluster سوم قرار می گیرند.



### سوال سوم

با الگوریتم agglomerative clustering و ماتریس فاصله ی زیر نقاط A تا D را (مرحله به مرحله) دسته بندی کنید و نمودار hierarchical آن ها را رسم کنید. فاصله ی دو دسته را نزدیک ترین نقاط آن ها در نظر بگیرید.

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

C			0	2
D				0

پاسخ

d	k	دسته ها	توضیح
0	4	{A}, {B}, {C}, {D}	در ابتدا هر نقطه یک دسته است.
1	3	{A, B}, {C}, {D}	چون دو دسته ی A و B نزدیک ترین دسته ها هستند و فاصله ی آن ها یک است با هم ترکیب می شوند.
2	2	{A, B}, {C, D}	چون دو دسته ی C و D نزدیک ترین دسته ها هستند و فاصله ی آن ها دو است با هم ترکیب می شوند.
3	1	{A, B, C, D}	در آخر دو دسته ی باقی مانده که فاصله ی آن ها سه است ترکیب می شوند و یک دسته باقی می ماند.

