Assignment 3 - written part Mohamad Mahdi Samadi 810101465

Q1

a) $K = 3$ cluster points = $A_1 \begin{vmatrix} 2 \\ 10 \end{vmatrix}$ ① $A_4 \begin{vmatrix} 5 \\ 8 \end{vmatrix}$ ② $A_7 \begin{vmatrix} 1 \\ 2 \end{vmatrix}$ ③

for each point we'll calculate its Euclidian distance from the K clusters and assign it to the nearest cluster

$A_2: \begin{vmatrix} 2 \\ 5 \end{vmatrix}$ $A_1 \to A_2 = 5$ $A_4 \to A_2 = 3\sqrt{2}$ $A_7 \to A_2 = \sqrt{10}$ ③

$A_3: \begin{vmatrix} 8 \\ 4 \end{vmatrix}$ $A_1 \to A_3 = 6\sqrt{2}$ $A_4 \to A_3 = 5$ $A_7 \to A_3 = \sqrt{53}$ ②

$A_5: \begin{vmatrix} 7 \\ 5 \end{vmatrix}$ $A_1 \to A_5 = 5\sqrt{2}$ $A_4 \to A_5 = \sqrt{13}$ $A_7 \to A_5 = 3\sqrt{5}$ ②

$A_6: \begin{vmatrix} 6 \\ 4 \end{vmatrix}$ $A_1 \to A_6 = 2\sqrt{13}$ $A_4 \to A_6 = \sqrt{17}$ $A_7 \to A_6 = \sqrt{29}$ ②

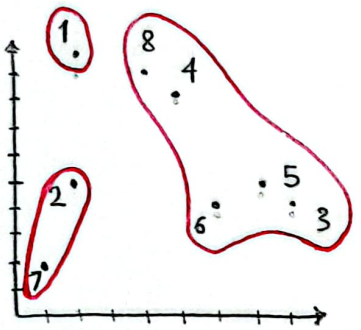
$A_8: \begin{vmatrix} 4 \\ 9 \end{vmatrix}$ $A_1 \to A_8 = \sqrt{5}$ $A_4 \to A_8 = \sqrt{2}$ $A_7 \to A_8 = \sqrt{58}$ ②

cluster ① = $\{A_1\}$ cluster ② = $\{A_3, A_4, A_5, A_6, A_8\}$ cluster ③ = $\{A_2, A_4\}$

b) new initial point of each cluster is average of all points in it

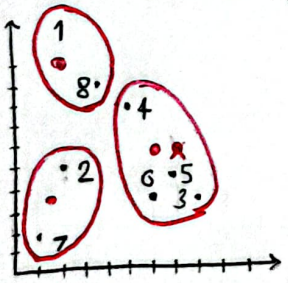
① : $\frac{A_1}{1} = \begin{vmatrix} 2 \\ 10 \end{vmatrix}$ ② : $\frac{A_3+A_4+A_5+A_6+A_8}{5} = \begin{vmatrix} 6 \\ 6 \end{vmatrix}$ ③ : $\frac{A_2+A_7}{2} = \begin{vmatrix} 1.5 \\ 3.5 \end{vmatrix}$

c)



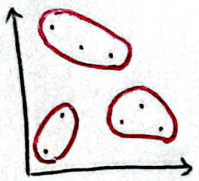
d)

Second epoch: ① : $\{A_1, A_8\}$ (core: $\begin{vmatrix} 6 \\ 6 \end{vmatrix}$) ② : $\{A_4, A_6, A_5, A_3\}$ ③ : $\{A_2, A_7\}$ core: $\begin{vmatrix} 3.5 \\ 6.5 \end{vmatrix}$



new cores: ① = $\frac{A_1+A_8}{2} = \begin{vmatrix} 3 \\ 9.5 \end{vmatrix}$ ② = $\frac{A_3+A_4+A_5+A_6}{4} = \begin{vmatrix} 6.5 \\ 5.25 \end{vmatrix}$ ③ = $\frac{A_2+A_7}{2} = \begin{vmatrix} 1.5 \\ 3.5 \end{vmatrix}$

third epoch: ① : $\{A_1, A_4, A_8\}$ ② : $\{A_6, A_5, A_3\}$ ③ : $\{A_2, A_7\}$



after third iteration the algorithm converges to a reasonable answer.

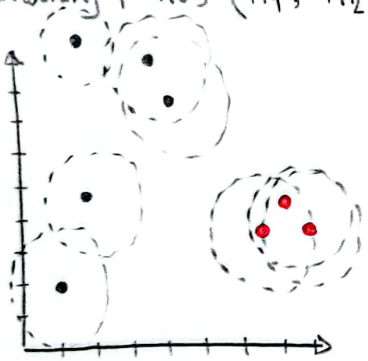
Assignment 3 - written part Mohamad Mahdi Samadi 81010146S

Q2) $A_1 \begin{vmatrix}2\\10\end{vmatrix}$ $A_2 \begin{vmatrix}2\\5\end{vmatrix}$ $A_3 \begin{vmatrix}8\\4\end{vmatrix}$ $A_4 \begin{vmatrix}5\\8\end{vmatrix}$ $A_5 \begin{vmatrix}7\\5\end{vmatrix}$ $A_6 \begin{vmatrix}6\\4\end{vmatrix}$ $A_7 \begin{vmatrix}1\\2\end{vmatrix}$ $A_8 \begin{vmatrix}4\\9\end{vmatrix}$

a) $r = 2$ and $n = 2$

point	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
Close points to it	X	X	A_5, A_6	A_8	A_3, A_6	A_3, A_5	X	A_4
is core?	NO	NO	Yes	NO	Yes	Yes	NO	NO

so we have 3 core points (A_3, A_5, A_6). we'll randomly select one of them (A_5). A_5 is close to A_3 and A_6. But there is no point close to any of these cores. So they'll form our first cluster. Since there are no other core points left, the DBSCAN algorithm finishes here and the remaining points (A_1, A_2, A_4, A_7, A_8) are considered noisy points.



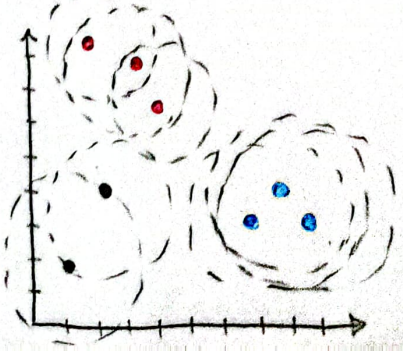
first cluster = $\{A_3, A_5, A_6\}$

noisy points = $\{A_1, A_2, A_4, A_7, A_8\}$

b) $r = \sqrt{10}$ and $n = 2$

point	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
Close points to it	A_8	A_7	A_5, A_6	A_8	A_3, A_6	A_3, A_5	A_2	A_1, A_4
is core?	NO	NO	Yes	NO	Yes	Yes	NO	Yes

core points: A_3, A_5, A_6, A_8. Let's choose a random core point and start the algorithm (A_8) There are zero core points and two non-core points (A_1, A_4) in proximity to A_8. Therefore the three of them form a cluster. A_3 is chosen as the next core point. There are two core points and zero non-core points near it. Since there are no core or non-core points in proximity to A_5 and A_6, these three points form our second cluster. Now that no core points remain, the algorithm is completed and the rest of the points (A_2, A_7) are considered noisy points.



first cluster = $\{A_1, A_4, A_8\}$

second cluster = $\{A_3, A_5, A_6\}$

noisy points = $\{A_2, A_7\}$

Assignment 3 - written part Mohamad Mahdi Samadi 810101465

Q_3)

$A \to B : 1$ $A \to C : 4$ $A \to D : 5$ $B \to C : 3$ $B \to D : 6$ $C \to D : 2$

Let's start the agglomerative algorithm. At the first step, we have 4 points and no clusters yet. The nearest points to each other are A and B, so we'll merge them to create our first cluster. Now two closest points or clusters to each other are C and D. Now we have two cluster left and the only choice is to merge them. The final result is A-B-C-D.

here is the step by step solution

1. [A] - [B] - [C] - [D] the min distance between any pair of points is 1 ($A \to B$)
2. [A B] - [C] - [D] the min distance between any pair of point/cluster is 2 ($C \to D$)
3. [A B] - [C D] the only possible move is to merge two clusters
4. [A B C D]