CSCI-GA.2560-001, Artificial Intelligence April 25, 2022

Solutions to Problem 1 of Homework 11 (5 Points)

Name: Anav Prasad (ap7152) *Due: 5PM on Monday, April 25*

Collaborators:

	A	B	C	D	E	F	G	H	I	J
x	352	933	192	293	512	444	4	1077	1033	701
y	768	1093	539	422	858	23	9	380	905	639

For all questions, using Euclidean distance squared for p, q of $= sqrt\left((p_x - q_x)^2 + (p_y - q_y)^2\right)$.

Questions

1. [5 pts.] Consider the above data.
    1. Trace the behavior of Hierarchical Clustering down to 3 clusters using single-linkage.
    2. Trace the behavior of Hierarchical Clustering down to 3 clusters using complete-linkage.

Solution:
Let's first of all consider the distance from every point to every other point:

	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
A	0.0	665.7	279.4	351.0	183.6	750.7	835.0	822.3	694.6	372.1
B	665.7	0.0	925.2	927.3	482.1	1176.4	1427.6	727.4	212.9	509.8
C	279.4	925.2	0.0	154.6	451.8	574.2	562.4	899.2	917.2	518.7
D	351.0	927.3	154.6	0.0	487.9	426.6	504.1	785.1	883.7	462.1
E	183.6	482.1	451.8	487.9	0.0	837.8	989.4	740.1	523.1	289.3
F	750.7	1176.4	574.2	426.6	837.8	0.0	440.2	726.7	1060.6	667.5
G	835.0	1427.6	562.4	504.1	989.4	440.2	0.0	1135.3	1364.4	939.5
H	822.3	727.4	899.2	785.1	740.1	726.7	1135.3	0.0	526.8	456.6
I	694.6	212.9	917.2	883.7	523.1	1060.6	1364.4	526.8	0.0	425.4
J	372.1	509.8	518.7	462.1	289.3	667.5	939.5	456.6	425.4	0.0

1. ***Single-Linkage:***

    - Step 1:
    Current Clusters: $\{A\}, \{B\}, \{C\}, \{D\}, \{E\}, \{F\}, \{G\}, \{H\}, \{I\}, \{J\}$
    Distances computed between clusters as per Single-Linkage:

	{A}	{B}	{C}	{D}	{E}	{F}	{G}	{H}	{I}	{J}
{A}	0.0	665.7	279.4	351.0	183.6	750.7	835.0	822.3	694.6	372.1
{B}		0.0	925.2	927.3	482.1	1176.4	1427.6	727.4	212.9	509.8
{C}			0.0	154.6	451.8	574.2	562.4	899.2	917.2	518.7
{D}				0.0	487.9	426.6	504.1	785.1	883.7	462.1
{E}					0.0	837.8	989.4	740.1	523.1	289.3
{F}						0.0	440.2	726.7	1060.6	667.5
{G}							0.0	1135.3	1364.4	939.5
{H}								0.0	526.8	456.6
{I}									0.0	425.4
{J}										0.0

The most minimum distance between clusters is between the clusters {C} and {D} i.e. 154.6. As such, they will be merged now.

- Step 2:
  Current Clusters: $\{A\}, \{B\}, \{C, D\}, \{E\}, \{F\}, \{G\}, \{H\}, \{I\}, \{J\}$
  Distances computed between clusters as per Single-Linkage:

	{A}	{B}	{C, D}	{E}	{F}	{G}	{H}	{I}	{J}
{A}	0.0	665.7	279.4	183.6	750.7	835.0	822.3	694.6	372.1
{B}		0.0	925.2	482.1	1176.4	1427.6	727.4	212.9	509.8
{C,D}			0.0	451.8	426.6	504.1	785.1	883.7	462.1
{E}				0.0	837.8	989.4	740.1	523.1	289.3
{F}					0.0	440.2	726.7	1060.6	667.5
{G}						0.0	1135.3	1364.4	939.5
{H}							0.0	526.8	456.6
{I}								0.0	425.4
{J}									0.0

The most minimum distance between clusters is between the clusters {A} and {E} i.e. 183.6. As such, they will be merged now.

- Step 3:
  Current Clusters: $\{A, E\}, \{B\}, \{C, D\}, \{F\}, \{G\}, \{H\}, \{I\}, \{J\}$
  Distances computed between clusters as per Single-Linkage:

	{A, E}	{B}	{C, D}	{F}	{G}	{H}	{I}	{J}
{A, E}	0.0	482.1	279.4	750.7	835.0	740.1	523.1	289.3
{B}		0.0	925.2	1176.4	1427.6	727.4	212.9	509.8
{C,D}			0.0	426.6	504.1	785.1	883.7	462.1
{F}				0.0	440.2	726.7	1060.6	667.5
{G}					0.0	1135.3	1364.4	939.5
{H}						0.0	526.8	456.6
{I}							0.0	425.4
{J}								0.0

The most minimum distance between clusters is between the clusters {B} and {I} i.e. 212.9. As such, they will be merged now.

- Step 4:
  Current Clusters: $\{A, E\}, \{B, I\}, \{C, D\}, \{F\}, \{G\}, \{H\}, \{J\}$
  Distances computed between clusters as per Single-Linkage:

	{A, E}	{B, I}	{C, D}	{F}	{G}	{H}	{J}
{A, E}	0.0	482.1	279.4	750.7	835.0	740.1	289.3
{B, I}		0.0	883.7	1060.6	1364.4	526.8	425.4
{C,D}			0.0	426.6	504.1	785.1	462.1
{F}				0.0	440.2	726.7	667.5
{G}					0.0	1135.3	939.5
{H}						0.0	456.6
{J}							0.0

The most minimum distance between clusters is between the clusters {A, E} and {C, D} i.e. 279.4. As such, they will be merged now.

- Step 5:
  Current Clusters: $\{\{A, E\}, \{C, D\}\}, \{B, I\}, \{F\}, \{G\}, \{H\}, \{J\}$
  Distances computed between clusters as per Single-Linkage:

	{{A,E}, {C, D}}	{B, I}	{F}	{G}	{H}	{J}
{{A,E}, {C, D}}	0.0	482.1	426.6	504.1	740.1	289.3
{B, I}		0.0	1060.6	1364.4	526.8	425.4
{F}			0.0	440.2	726.7	667.5
{G}				0.0	1135.3	939.5
{H}					0.0	456.6
{J}						0.0

The most minimum distance between clusters is between the clusters {{A,E}, {C, D}} and {J} i.e. 289.3. As such, they will be merged now.

- Step 6:
  Current Clusters: $\{\{\{A, E\}, \{C, D\}\}, \{J\}\}, \{B, I\}, \{F\}, \{G\}, \{H\}$
  Distances computed between clusters as per Single-Linkage:

	{{{A,E}, {C, D}}, {J}}	**{B, I}**	**{F}**	**{G}**	**{H}**
{{{A,E}, {C, D}}, {J}}	0.0	425.4	426.6	504.1	456.6
{B, I}		0.0	1060.6	1364.4	526.8
{F}			0.0	440.2	726.7
{G}				0.0	1135.3
{H}					0.0

  The most minimum distance between clusters is between the clusters {{{A,E}, {C, D}}, {J}} and {B, I} i.e. 425.4. As such, they will be merged now.

- Step 7:
  Current Clusters: $\{\{\{\{A, E\}, \{C, D\}\}, \{J\}\}, \{B, I\}\}, \{F\}, \{G\}, \{H\}$
  Let C_1 represent {{{{A,E}, {C, D}}, {J}}, {B, I}} now.
  Distances computed between clusters as per Single-Linkage:

	C_1	**{F}**	**{G}**	**{H}**
C_1	0.0	426.6	504.1	456.6
{F}		0.0	440.2	726.7
{G}			0.0	1135.3
{H}				0.0

  The most minimum distance between clusters is between the clusters C_1 and {F} i.e. 426.6. As such, they will be merged now.

- Step 8:
  Current Clusters: $\{\{\{\{\{A, E\}, \{C, D\}\}, \{J\}\}, \{B, I\}\}, \{F\}\}, \{G\}, \{H\}$
  Let C_2 represent {{{{{A,E}, {C, D}}, {J}}, {B, I}}, {F}} now.
  Distances computed between clusters as per Single-Linkage:

	C_2	**{G}**	**{H}**
C_2	0.0	440.2	456.6
{G}		0.0	1135.3
{H}			0.0

  Since we have reached the target of 3 clusters, the process will stop now.

Therefore, the 3 clusters obtained are as follows:
{{{{{A,E}, {C, D}}, {J}}, {B, I}}, {F}}, {G}, {H}

2. ***Complete-Linkage:***

- Step 1:
  Current Clusters: $\{A\}, \{B\}, \{C\}, \{D\}, \{E\}, \{F\}, \{G\}, \{H\}, \{I\}, \{J\}$
  Distances computed between clusters as per Single-Linkage:

	{A}	**{B}**	**{C}**	**{D}**	**{E}**	**{F}**	**{G}**	**{H}**	**{I}**	**{J}**
{A}	0.0	665.7	279.4	351.0	183.6	750.7	835.0	822.3	694.6	372.1
{B}		0.0	925.2	927.3	482.1	1176.4	1427.6	727.4	212.9	509.8
{C}			0.0	154.6	451.8	574.2	562.4	899.2	917.2	518.7
{D}				0.0	487.9	426.6	504.1	785.1	883.7	462.1
{E}					0.0	837.8	989.4	740.1	523.1	289.3
{F}						0.0	440.2	726.7	1060.6	667.5
{G}							0.0	1135.3	1364.4	939.5
{H}								0.0	526.8	456.6
{I}									0.0	425.4
{J}										0.0

  The most minimum distance between clusters is between the clusters {C} and {D} i.e. 154.6. As such, they will be merged now.

- Step 2:
  Current Clusters: $\{A\}, \{B\}, \{C, D\}, \{E\}, \{F\}, \{G\}, \{H\}, \{I\}, \{J\}$
  Distances computed between clusters as per Single-Linkage:

	{A}	**{B}**	**{C, D}**	**{E}**	**{F}**	**{G}**	**{H}**	**{I}**	**{J}**
{A}	0.0	665.7	351.0	183.6	750.7	835.0	822.3	694.6	372.1
{B}		0.0	927.3	482.1	1176.4	1427.6	727.4	212.9	509.8
{C, D}			0.0	487.9	574.2	562.4	899.2	917.2	518.7
{E}				0.0	837.8	989.4	740.1	523.1	289.3
{F}					0.0	440.2	726.7	1060.6	667.5
{G}						0.0	1135.3	1364.4	939.5
{H}							0.0	526.8	456.6
{I}								0.0	425.4
{J}									0.0

  The most minimum distance between clusters is between the clusters {A} and {E} i.e. 183.6. As such, they will be merged now.

- Step 3:
  Current Clusters: $\{A, E\}, \{B\}, \{C, D\}, \{F\}, \{G\}, \{H\}, \{I\}, \{J\}$
  Distances computed between clusters as per Single-Linkage:

	{A, E}	**{B}**	**{C, D}**	**{F}**	**{G}**	**{H}**	**{I}**	**{J}**
{A, E}	0.0	665.7	487.9	837.8	989.4	822.3	694.6	372.1
{B}		0.0	927.3	1176.4	1427.6	727.4	212.9	509.8
{C, D}			0.0	574.2	562.4	899.2	917.2	518.7
{F}				0.0	440.2	726.7	1060.6	667.5
{G}					0.0	1135.3	1364.4	939.5
{H}						0.0	526.8	456.6
{I}							0.0	425.4
{J}								0.0

The most minimum distance between clusters is between the clusters {B} and {I} i.e. 212.9. As such, they will be merged now.

- Step 4:
  Current Clusters: $\{A, E\}, \{B, I\}, \{C, D\}, \{F\}, \{G\}, \{H\}, \{J\}$
  Distances computed between clusters as per Single-Linkage:

	{A, E}	**{B, I}**	**{C, D}**	**{F}**	**{G}**	**{H}**	**{J}**
{A, E}	0.0	694.6	487.9	837.8	989.4	822.3	372.1
{B, I}		0.0	927.3	1176.4	1427.6	727.4	509.8
{C, D}			0.0	574.2	562.4	899.2	518.7
{F}				0.0	440.2	726.7	667.5
{G}					0.0	1135.3	939.5
{H}						0.0	456.6
{J}							0.0

The most minimum distance between clusters is between the clusters {A, E} and {J} i.e. 372.1. As such, they will be merged now.

- Step 5:
  Current Clusters: $\{\{A, E\}, \{J\}\}, \{B, I\}, \{C, D\}, \{F\}, \{G\}, \{H\}$
  Distances computed between clusters as per Single-Linkage:

	{{A, E}, {J}}	**{B, I}**	**{C, D}**	**{F}**	**{G}**	**{H}**
{{A, E}, {J}}	0.0	694.6	518.7	837.8	989.4	822.3
{B, I}		0.0	927.3	1176.4	1427.6	727.4
{C, D}			0.0	574.2	562.4	899.2
{F}				0.0	440.2	726.7
{G}					0.0	1135.3
{H}						0.0

The most minimum distance between clusters is between the clusters {F} and {G} i.e. 440.2. As such, they will be merged now.

- Step 6:
  Current Clusters: $\{\{A, E\}, \{J\}\}, \{B, I\}, \{C, D\}, \{F, G\}, \{H\}$
  Distances computed between clusters as per Single-Linkage:

	{{A, E}, {J}}	**{B, I}**	**{C, D}**	**{F, G}**	**{H}**
{{A, E}, {J}}	0.0	694.6	518.7	989.4	822.3
{B, I}		0.0	927.3	1427.6	727.4
{C, D}			0.0	574.2	899.2
{F, G}				0.0	1135.3
{H}					0.0

  The most minimum distance between clusters is between the clusters {{A, E}, {J}} and {C, D} i.e. 518.7. As such, they will be merged now.

- Step 7:
  Current Clusters: $\{\{\{A, E\}, \{J\}\}, \{C, D\}\}, \{B, I\}, \{F, G\}, \{H\}$
  Distances computed between clusters as per Single-Linkage:

	{{{A, E}, {J}}, {C, D}}	**{B, I}**	**{F, G}**	**{H}**
{{{A, E}, {J}}, {C, D}}	0.0	927.3	989.4	899.2
{B, I}		0.0	1427.6	727.4
{F, G}			0.0	1135.3
{H}				0.0

  The most minimum distance between clusters is between the clusters {B, I} and {H} i.e. 727.4. As such, they will be merged now.

- Step 8:
  Current Clusters: $\{\{\{A, E\}, \{J\}\}, \{C, D\}\}, \{\{B, I\}, \{H\}\}, \{F, G\}$
  Distances computed between clusters as per Single-Linkage:

	{{{A, E}, {J}}, {C, D}}	**{{B, I}, {H}}**	**{F, G}**
{{{A, E}, {J}}, {C, D}}	0.0	927.3	989.4
{{B, I}, {H}}		0.0	1427.6
{F, G}			0.0

  Since we have reached the target of 3 clusters, the process will stop now.

Therefore, the 3 clusters obtained are as follows:
{{{A, E}, {J}}, {C, D}}, {{B, I}, {H}}, {F, G}

□

CSCI-GA.2560-001, Artificial Intelligence April 25, 2022

Solutions to Problem 2 of Homework 11 (5 Points)

Name: Anav Prasad (ap7152) *Due: 5PM on Monday, April 25*

Collaborators:

2. [5 pts.] Trace the behavior of k-Means, using 2-decimal places for new centroids, with:
   1. $k = 3$, initial centroids $(500, 10), (200, 700), (800, 200)$
   2. $k = 3$, initial centroids of A, B, and C

Solution:

1. **PART 1:** $k = 3$, initial centroids $(500, 10), (200, 700), (800, 200)$

   - Step 1:
     Current Centroids: $(500, 10)$, $(200, 700)$, $(800, 200)$
     Distances:

	(500, 10)	**(200, 700)**	**(800, 200)**
A	772.3	166.5	723.4
B	1166.4	831.7	902.8
C	612.1	161.2	696.1
D	461.1	293.1	553.5
E	848.1	349.7	718.3
F	57.5	719.6	397.6
G	496.0	718.3	818.6
H	685.4	933.6	330.3
I	1041.7	857.9	742.5
J	660.3	504.7	450.0

     Therefore, the updated clusters are:
     - $(500, 10)$: {F, G}
       Updated centroid:

$$
\begin{aligned}
&= \frac{((444, 23) + (4, 9))}{2} \\
&= \left(\frac{444 + 4}{2}, \frac{23 + 9}{2}\right) \\
&= (224.00, 16.00)
\end{aligned}
$$

- $(200, 700)$: {A, B, C, D, E}
  Updated centroid:

$$= \frac{((352, 768) + (933, 1093) + (192, 539) + (293, 422) + (512, 858))}{5}$$
$$= \left(\frac{352 + 933 + 192 + 293 + 512}{5}, \frac{768 + 1093 + 539 + 422 + 858}{5}\right)$$
$$= (456.40, 736.00)$$

- $(800, 200)$: {H, I, J}
  Updated centroid:

$$= \frac{((1077, 380) + (1033, 905) + (701, 639))}{3}$$
$$= \left(\frac{1077 + 1033 + 701}{3}, \frac{380 + 905 + 639}{3}\right)$$
$$= (937.00, 641.33)$$

- Step 2:
  Current Centroids: $(224.00, 16.00)$, $(456.40, 736.00)$, $(937.00, 641.33)$
  Distances:

	(224.00, 16.00)	**(456.40,736.00)**	**(937.00,641.33)**
A	762.8	109.2	598.6
B	1289.4	595.5	451.7
C	524.0	329.7	752.0
D	411.8	354.0	680.3
E	889.9	134.1	477.0
F	220.1	713.1	790.8
G	220.1	856.3	1127.1
H	927.4	715.5	296.5
I	1202.0	600.9	280.6
J	784.6	263.1	236.0

  Therefore, the updated clusters are:
  - $(224.00, 16.00)$: {F, G}
    No Update
  - $(456.40, 736.00)$: {A, C, D, E}

Updated centroid:

$$= \frac{((352, 768) + (192, 539) + (293, 422) + (512, 858))}{4}$$
$$= \left(\frac{352 + 192 + 293 + 512}{4}, \frac{768 + 539 + 422 + 858}{4}\right)$$
$$= (337.25, 646.75)$$

- $(937.00, 641.33)$: {B, H, I, J}
  Updated centroid:

$$= \frac{((933, 1093) + (1077, 380) + (1033, 905) + (701, 639))}{4}$$
$$= \left(\frac{933 + 1077 + 1033 + 701}{4}, \frac{1093 + 380 + 905 + 639}{4}\right)$$
$$= (936.00, 754.25)$$

- Step 3:
  Current Centroids: $(224.00, 16.00)$, $(337.25, 646.75)$, $(936.00, 754.25)$
  Distances:

	(224.00, 16.00)	**(456.40,736.00)**	**(937.00,641.33)**
A	762.8	122.1	584.2
B	1289.4	744.4	338.8
C	524.0	180.9	774.5
D	411.8	229.1	723.8
E	889.9	274.2	436.5
F	220.1	632.8	881.4
G	220.1	719.6	1193.3
H	927.4	786.4	399.9
I	1202.0	742.1	179.3
J	784.6	363.8	261.7

  Therefore, the updated clusters are:
  - $(224.00, 16.00)$: {F, G}
    No Update
  - $(456.40, 736.00)$: {A, C, D, E}
    No Update
  - $(937.00, 641.33)$: {B, H, I, J}
    No Update

Therefore, the final cluster obtained is:
{F, G}, {A, C, D, E}, {B, H, I, J}

2. **PART 2:** $k = 3$, initial centroids of A, B, and C

- Step 1:
  Current Centroids: $(352.00, 768.00)$, $(933.00, 1093.00)$, $(192.00, 539.00)$
  Distances:

	(352.00, 768.00)	**(933.00, 1093.00)**	**(192.00, 539.00)**
A	0.0	665.7	279.4
B	665.7	0.0	925.2
C	279.4	925.2	0.0
D	351.0	927.3	154.6
E	183.6	482.1	451.8
F	750.7	1176.4	574.2
G	835.0	1427.6	562.4
H	822.3	727.4	899.2
I	694.6	212.9	917.2
J	372.1	509.8	518.7

  Therefore, the updated clusters are:
  - $(352.00, 768.00)$: {A, E, J}
    Updated centroid:

$$\begin{aligned} &= \frac{((352, 768) + (512, 858) + (701, 639))}{3} \\ &= \left(\frac{352 + 512 + 701}{3}, \frac{768 + 858 + 639}{3}\right) \\ &= (521.67, 755.00) \end{aligned}$$

  - $(933.00, 1093.00)$: {B, H, I}
    Updated centroid:

$$\begin{aligned} &= \frac{((933, 1093) + (1077, 380) + (1033, 905))}{3} \\ &= \left(\frac{933 + 1077 + 1033}{3}, \frac{1093 + 380 + 905}{3}\right) \\ &= (1014.33, 792.67) \end{aligned}$$

  - $(192.00, 539.00)$: {C, D, F, G}

Updated centroid:

$$
\begin{aligned}
&= \frac{((192, 539) + (293, 422) + (444, 23) + (4, 9))}{4} \\
&= \left(\frac{192 + 293 + 444 + 4}{4}, \frac{539 + 422 + 23 + 9}{4}\right) \\
&= (233.25, 248.25)
\end{aligned}
$$

- Step 2:
  Current Centroids: $(521.67, 755.00)$, $(1014.33, 792.67)$, $(233.25, 248.25)$
  Distances:

	(521.67, 755.00)	**((1014.33, 792.67)**	**(233.25, 248.25)**
A	170.2	662.8	533.1
B	532.4	311.1	1096.9
C	394.1	860.6	293.7
D	404.0	811.0	183.7
E	103.5	506.6	670.4
F	736.1	958.0	308.5
G	908.0	1278.6	331.4
H	670.1	417.4	854.0
I	532.9	113.9	1034.9
J	213.6	349.0	609.5

  Therefore, the updated clusters are:
  - $(521.67, 755.00)$: {A, E, J}
    No Update
  - $(1014.33, 792.67)$: {B, H, I}
    No Update
  - $(233.25, 248.25)$: {C, D, F, G}
    No Update

Therefore, the final cluster obtained is:
{A, E, J}, {B, H, I}, {C, D, F, G}

□

CSCI-GA.2560-001, Artificial Intelligence April 25, 2022

Solutions to Problem 3 of Homework 11 (5 Points)

Name: Anav Prasad (ap7152) *Due: 5PM on Monday, April 25*

Collaborators:

1. Is k-means in the general case guaranteed to terminate? Why or why not?

   Solution:
   Yes, K-Means is guaranteed to converge. That is so because in every iteration, the sum of distances to the center is ensured to be reduced (by its definition). Furthermore, K-Means can be, for the purpose of proof of its convergence, reformulated into a Expectation-Minimization(EM) Algorithm and, every EM algorithm converges. □

2. True or False: linear functions are good candidates for neural network perceptrons? Explain your answer.

   Solution:
   False. Non-Linear functions must be used as activation functions in neural network perceptrons because a linear function would invalidate the presence of additional layers of the network and make it behave as if there is only one single layer there. That is to say, the network will not be able to learn non-linear relationships from the data. □

3. True or False: finding a linear separator for a training set is NP-hard? Explain your answer.

   Solution:
   True. Unless the training set is linearly separable (or P=NP), there is no polynomial time algorithm to find optimal parameters for the separator that minimizes the loss. In other words, the problem is NP-Hard. □

4. How is a support vector machine ("SVM") different than a linear separator? Explain.

   Solution:
   The difference is that SVMs use kernels to make non-separable data separable whereas linear separators are use lines (or hyperplanes) to try and separate the data, often in vain. □

5. What is the salient feature of Deep Learning as opposed to traditional Supervised Learning?

   Solution:
   The major difference is in the feature representation. In supervised learning, features are derived or abstract concepts/properties that are created explicitly. In deep learning, the same are automatically generated in a number of hidden layers.
   As a consequence, the number of weights in deep learning models are very large as compared to supervised learning models. □