**502571-3 2nd Trimester 2022/2023 HW#5**

# Topics: K-Means Algo Sections: 2233 & 3827

# Due Date: Monday 13/02/2023 – 2:00 pm

# \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Use the K-Means algorithm and Euclidean distance to cluster the following 8 instances into 3 clusters: I1 = (2, 10); I2 = (2, 5); I3 = (8, 4); I4 = (5, 8); I5 = (7, 5); I6 = (6, 4); I7 = (1, 2); I8 = (4, 9).

Suppose that the initial seeds (center of each cluster or centroid) are: I1, I4, and I7. Run the K-Means algorithm until the convergence (the algorithm is stopped).

|  |  |
| --- | --- |
| Attribute1 | Attribute 2 |
| 2 | 10 |
| 2 | 5 |
| 8 | 4 |
| 5 | 8 |
| 7 | 5 |
| 6 | 4 |
| 1 | 2 |
| 4 | 9 |

Step 1:

Initial choice of centroid

|  |  |  |
| --- | --- | --- |
| Centroid 1 | 2 | 10 |
| Centroid 2 | 5 | 8 |
| Centroid 3 | 1 | 2 |

Step 2:

Here I used Euclidean distance:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Attribute1 | Attribute 2 | Centroid 1 | Centroid 2 | Centroid 3 | Cluster |
| 2 | 10 | 0 | 3.60 | 8.06 | 1 |
| 2 | 5 | 5 | 4.24 | 3.16 | 3 |
| 8 | 4 | 8.49 | 5 | 7.28 | 2 |
| 5 | 8 | 3.60 | 0 | 7.21 | 2 |
| 7 | 5 | 7.07 | 3.60 | 6.70 | 2 |
| 6 | 4 | 7.21 | 4.12 | 5.39 | 2 |
| 1 | 2 | 8.06 | 7.21 | 0 | 3 |
| 4 | 9 | 2.23 | 1.41 | 7.62 | 2 |

Step 3:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Initial | | After first iteration | |
| x | y | x | y |
| Centroid 1 | 2 | 10 | 2 | 10 |
| Centroid 2 | 5 | 8 | 6 | 6 |
| Centroid 3 | 1 | 2 | 1.5 | 3.5 |

Step 4:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Attribute1 | Attribute 2 | Centroid 1 | Centroid 2 | Centroid 3 | Cluster |
| 2 | 10 | 0 | 5.65 | 6.51 | 1 |
| 2 | 5 | 5 | 4.12 | 1.58 | 3 |
| 8 | 4 | 8.49 | 2.28 | 6.51 | 2 |
| 5 | 8 | 3.60 | 2.23 | 5.70 | 2 |
| 7 | 5 | 7.07 | 1.41 | 5.70 | 2 |
| 6 | 4 | 7.21 | 2 | 4.52 | 2 |
| 1 | 2 | 8.06 | 6.40 | 1.58 | 3 |
| 4 | 9 | 2.23 | 3.60 | 6.04 | 1 |

Step5:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Initial | | After first iteration | | After second iteration | |
| x | y | x | y | x | y |
| Centroid 1 | 2 | 10 | 2 | 10 | 3 | 6.5 |
| Centroid 2 | 5 | 8 | 6 | 6 | 4 | 4 |
| Centroid 3 | 1 | 2 | 1.5 | 3.5 | 1.5 | 3.5 |

Step 6:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Attribute1 | Attribute 2 | Centroid 1 | Centroid 2 | Centroid 3 | Cluster |
| 2 | 10 | 3.64 | 6.23 | 6.51 | 1 |
| 2 | 5 | 1.80 | 2.23 | 1.58 | 3 |
| 8 | 4 | 5.59 | 4 | 6.51 | 2 |
| 5 | 8 | 2.5 | 4.12 | 5.70 | 1 |
| 7 | 5 | 4.27 | 3.16 | 5.70 | 2 |
| 6 | 4 | 3.90 | 2 | 4.52 | 2 |
| 1 | 2 | 4.92 | 3.60 | 1.58 | 3 |
| 4 | 9 | 2.69 | 5 | 6.04 | 1 |

Step7:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Initial | | After first iteration | | After second iteration | | After third iteration | |
| x | y | x | y | x | y | x | y |
| Centroid 1 | 2 | 10 | 2 | 10 | 3 | 6.5 | 3.66 | 9 |
| Centroid 2 | 5 | 8 | 6 | 6 | 4 | 4 | 7 | 4.33 |
| Centroid 3 | 1 | 2 | 1.5 | 3.5 | 1.5 | 3.5 | 1.5 | 3.5 |

Step8:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Attribute1 | Attribute 2 | Centroid 1 | Centroid 2 | Centroid 3 | Cluster |
| 2 | 10 | 1.93 | 7.55 | 6.51 | 1 |
| 2 | 5 | 4.33 | 5.04 | 1.58 | 3 |
| 8 | 4 | 6.62 | 1.05 | 6.51 | 2 |
| 5 | 8 | 1.67 | 4.17 | 5.70 | 1 |
| 7 | 5 | 5.21 | 0.67 | 5.70 | 2 |
| 6 | 4 | 5.52 | 1.05 | 4.52 | 2 |
| 1 | 2 | 7.48 | 6.43 | 1.58 | 3 |
| 4 | 9 | 0.34 | 5.55 | 6.04 | 1 |

Now as we can see these are the same clusters as in Step 6