Yuan Wang

440 Data Mining

Prof. Jiebo Luo

Homework 5

**10.2**

**(a)**

The initial clusters are based on the Euclidean distance between the rest points and the pre-assigned center A1, B1, C1. The minimum distances are labeled in red color:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | X | Y | A1 | B1 | C1 |
| A1 | 2 | 10 |  |  |  |
| A2 | 2 | 5 | 5.0 | 4.2 | 3.2 |
| A3 | 8 | 4 | 8.5 | 5.0 | 7.3 |
| B1 | 5 | 8 |  |  |  |
| B2 | 7 | 5 | 7.1 | 3.6 | 6.7 |
| B3 | 6 | 4 | 7.2 | 4.1 | 5.4 |
| C1 | 1 | 2 |  |  |  |
| C2 | 4 | 9 | 2.2 | 1.4 | 7.6 |

So the first round three clusters are:

{A1}

{B1, A3, B2, B3, C2}

{C1, A2}

**(b)**

The final three clusters are:

{A1, C2, B1}

{A2, C1}

{A3, B2, B3}

A picture containing text, application

Description automatically generated