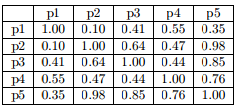
**Name: Nikhil Jadhav**

**Student ID: 801075504**

**Cloud Computing for Data Analysis**

**Exercise 10 : Hierarchical Clustering**

Use the similarity matrix in following table to perform single and complete link hierarchical clustering. Show your results by drawing a dendrogram. The dendrogram should clearly show the order in which the points are merged.



Step 1:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | p1 | p2 | p3 | p4 | p5 |
| p1 | 0 | - | - | - | - |
| p2 | 1-0.10 | 0 | - | - | - |
| p3 | 1-0.41 | 1-0.64 | 0 | - | - |
| p4 | 1-0.55 | 1-0.47 | 1-0.44 | 0 | - |
| p5 | 1-0.35 | 1-0.98 | 1-0.85 | 1-0.76 | 0 |

Step 2:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | p1 | p2 | p3 | p4 | p5 |
| p1 | 0 | - | - | - | - |
| p2 | 0.9 | 0 | - | - | - |
| p3 | 0.59 | 0.36 | 0 | - | - |
| p4 | 0.45 | 0.53 | 0.56 | 0 | - |
| p5 | 0.65 | 0.02 | 0.15 | 0.24 | 0 |

Step 3:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | p1 | (p2, p5) | p3 | p4 |
| p1 | 0 | - | - | - |
| (p2, p5) | 0.65 | 0 | - | - |
| p3 | 0.59 | 0.15 | 0 | - |
| p4 | 0.45 | 0.24 | 0.56 | 0 |

Step 4:

|  |  |  |  |
| --- | --- | --- | --- |
|  | p1 | ((p2, p5), p3) | p4 |
| p1 | 0 | - | - |
| ((p2, p5), p3) | 0.59 | 0 | - |
| p4 | 0.45 | 0.24 | 0 |

Step 5:

|  |  |  |
| --- | --- | --- |
|  | p1 | (((p2, p5), p3), p4) |
| p1 | 0 | - |
| (((p2, p5), p3), p4) | 0.45 | 0 |

So, final link is : ((((p2, p5), p3), p4), p1)

Dendrogram for Single Link:

2

1

4

3

5

0.55

0.65

0.75

0.85

0.95

1

Dendrogram for Complete Link:

0.1

0.2

0.3

2

4

1

3

5

0.5

0.6

0.7

0.8

0.9

1

0.4