CSE 469: Assignment 3

k-Means and Hierarchical Clustering

Peter M. VanNostrand

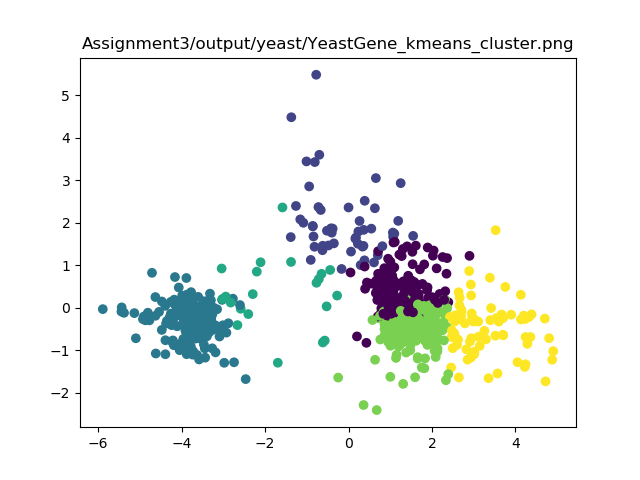
11/04/2019

# Assignment 3: Clustering

## k-Means Cluster Centroids: YeastGene Dataset

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| -0.2413 | -0.1288 | 0.0623 | 0.1734 | 0.2179 | 1.6517 | 1.9053 |
| -0.9535 | -1.4716 | 0.0775 | -0.1795 | -1.0048 | 1.1512 | 0.9688 |
| 0.1651 | 0.0917 | -0.1039 | -0.5526 | -0.6301 | -1.7232 | -1.7548 |
| 0.0233 | 0.2508 | -0.2770 | -0.3640 | -0.7354 | -0.8946 | 0.7001 |
| -0.0016 | 0.1565 | 0.3563 | 0.7016 | 1.0097 | 1.8423 | 1.6434 |
| -0.0393 | 0.1539 | 0.4361 | 1.1058 | 1.4487 | 3.0163 | 2.8294 |

## PCA of Clusters: YeastGene Dataset



## Hierarchical Clustering: Utilities

|  |  |  |  |
| --- | --- | --- | --- |
| Merging | Cluster1 | Cluster2 | New Cluster |
| 0 | 12 | 21 | 23 |
| 1 | 10 | 13 | 24 |
| 2 | 4 | 24 | 25 |
| 3 | 7 | 23 | 26 |
| 4 | 20 | 25 | 27 |
| 5 | 14 | 19 | 28 |
| 6 | 1 | 18 | 29 |
| 7 | 15 | 26 | 30 |
| 8 | 28 | 29 | 31 |
| 9 | 2 | 27 | 32 |
| 10 | 8 | 16 | 33 |
| 11 | 30 | 32 | 34 |
| 12 | 22 | 34 | 35 |
| 13 | 9 | 31 | 36 |
| 14 | 35 | 36 | 37 |
| 15 | 6 | 37 | 38 |
| 16 | 3 | 38 | 39 |
| 17 | 33 | 39 | 40 |
| 18 | 17 | 40 | 41 |
| 19 | 11 | 41 | 42 |
| 20 | 5 | 42 | 43 |

## Code: k-Means Clustering

### Assign Cluster



### Get Centroid



## Code: Hierarchical Clustering

### Merge Cluster



### Update Distance

