

Assignment 2 - Clustering

Part 1: Theory

None for this Assignment

Part 2: Clustering

(a) The initial setup visualization

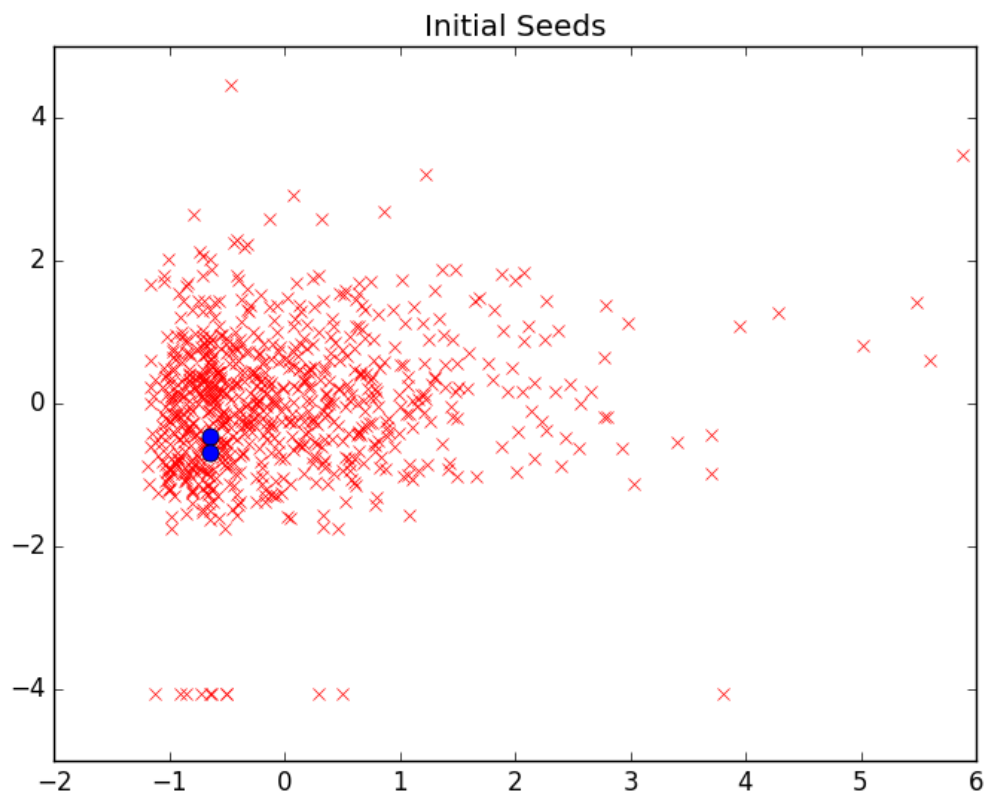


Figure 1. Initial data seeded with 2 reference vectors

(b) The initial cluster assignment visualization

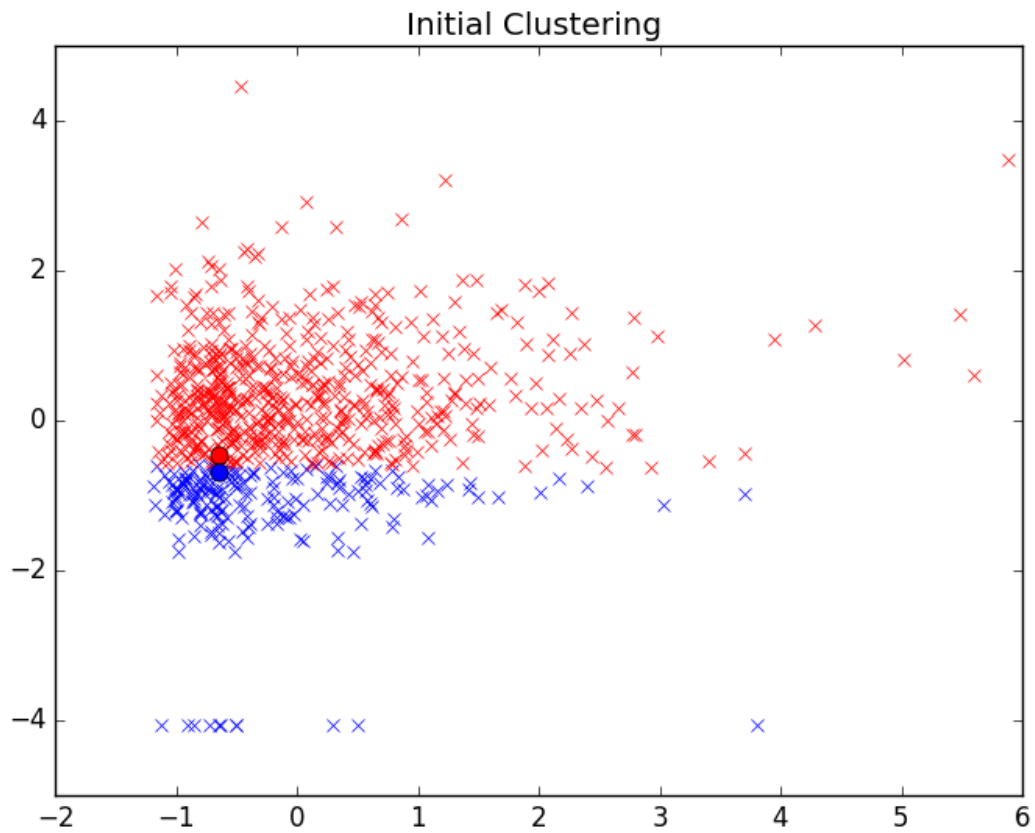


Figure 2. 1st iteration of k-means clustering given seeded reference vectors

(c) The final cluster assignment visualization

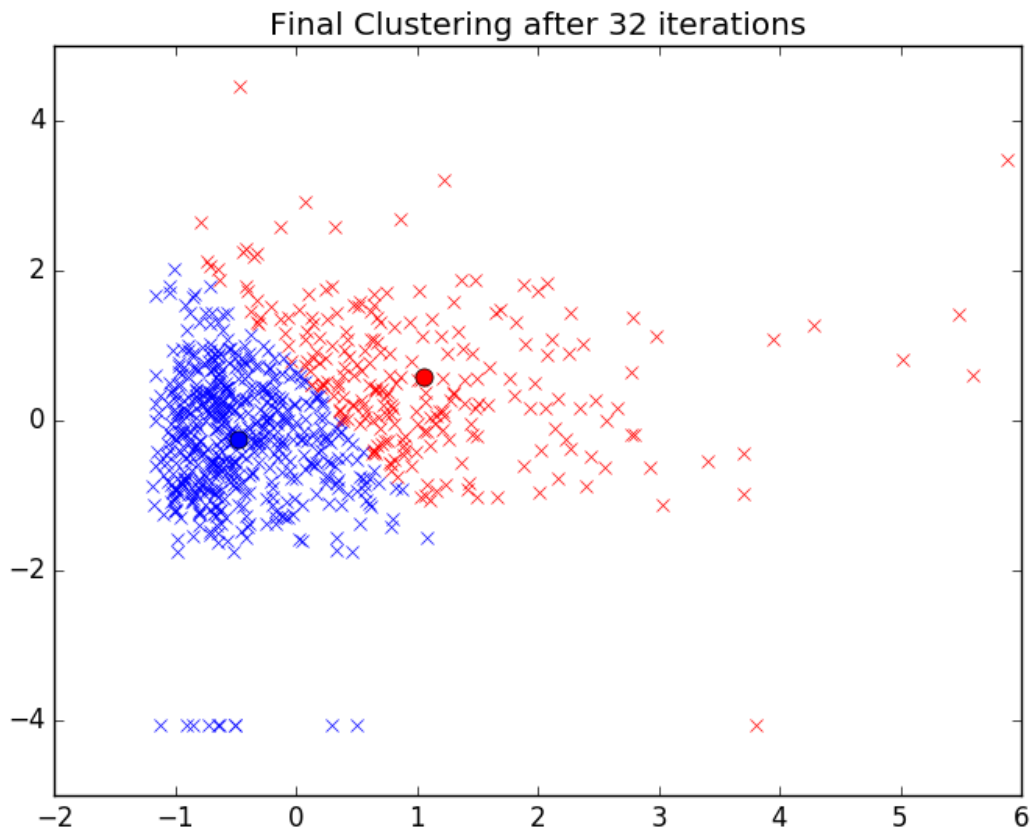


Figure 3. Final k-means clustering after converging to minimum distance threshold between clustering iterations

Part 3: Purity Value

Purity of Cluster 0:

0.533613445378

Purity of Cluster 1:

0.733962264151

Total Purity:

0.671875

(Purity of clusters computed from python file: purity.py)

Part 4: Clustering Part 2

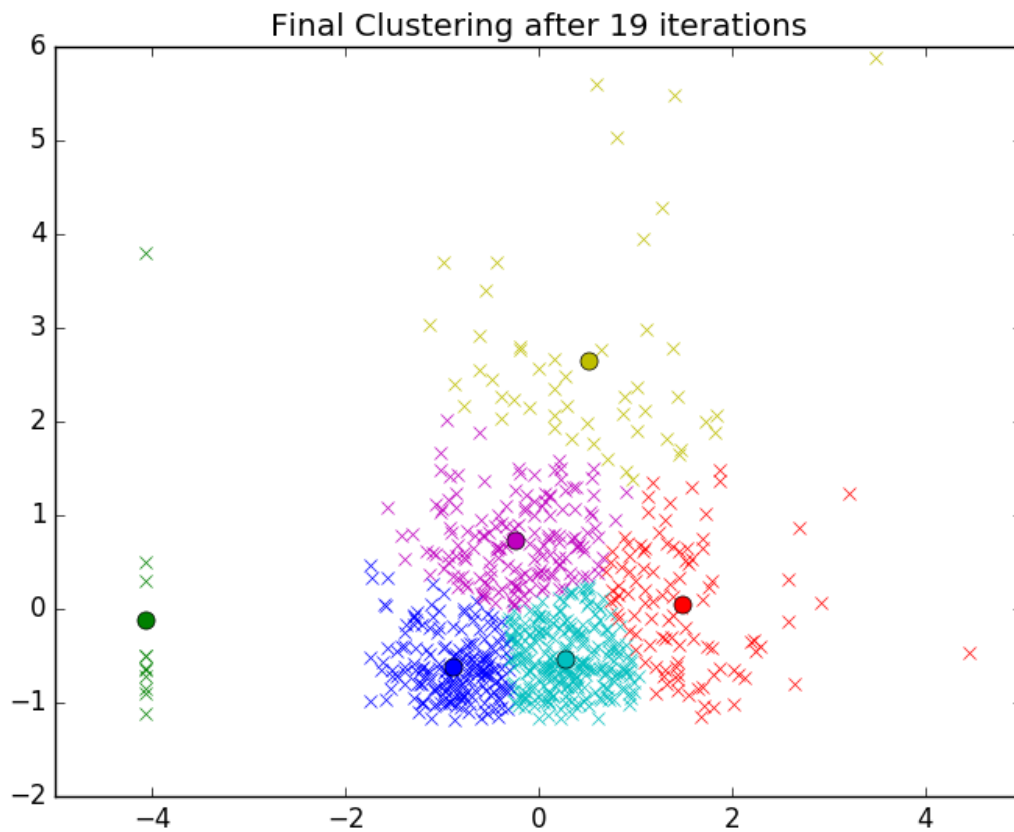


Figure 4. Features 6,7, $k=6$, plot features 6, 7

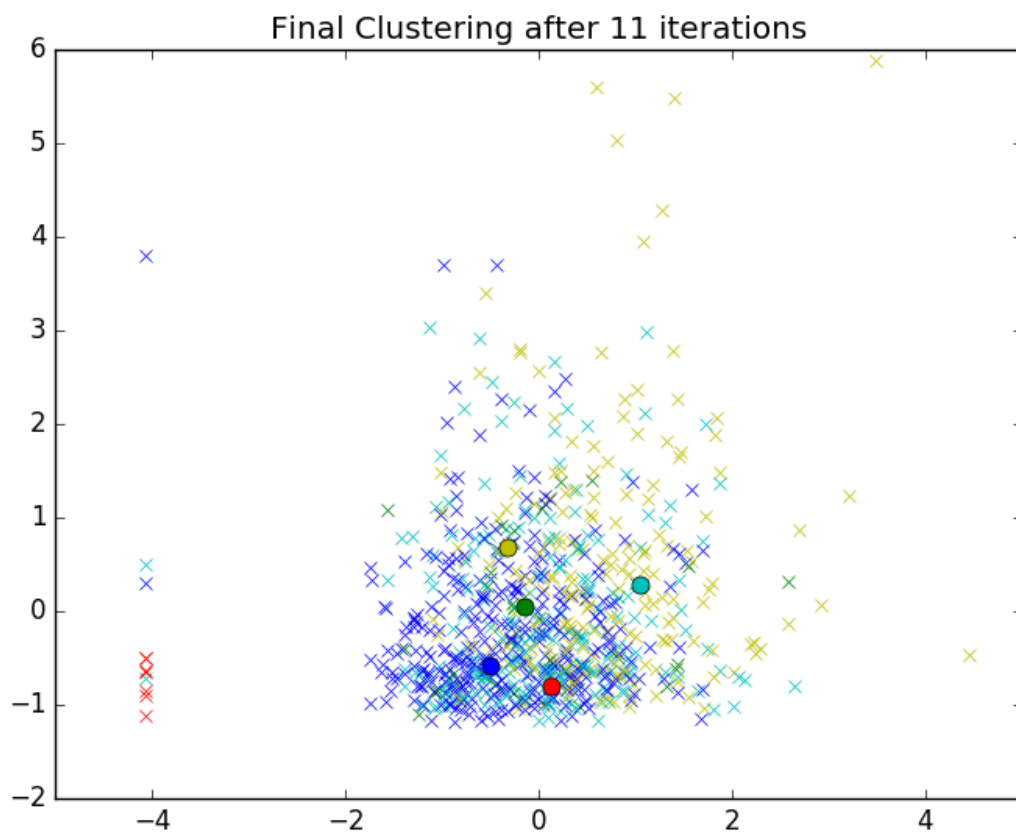


Figure 5. All features, $k=5$, plot features 6, 7

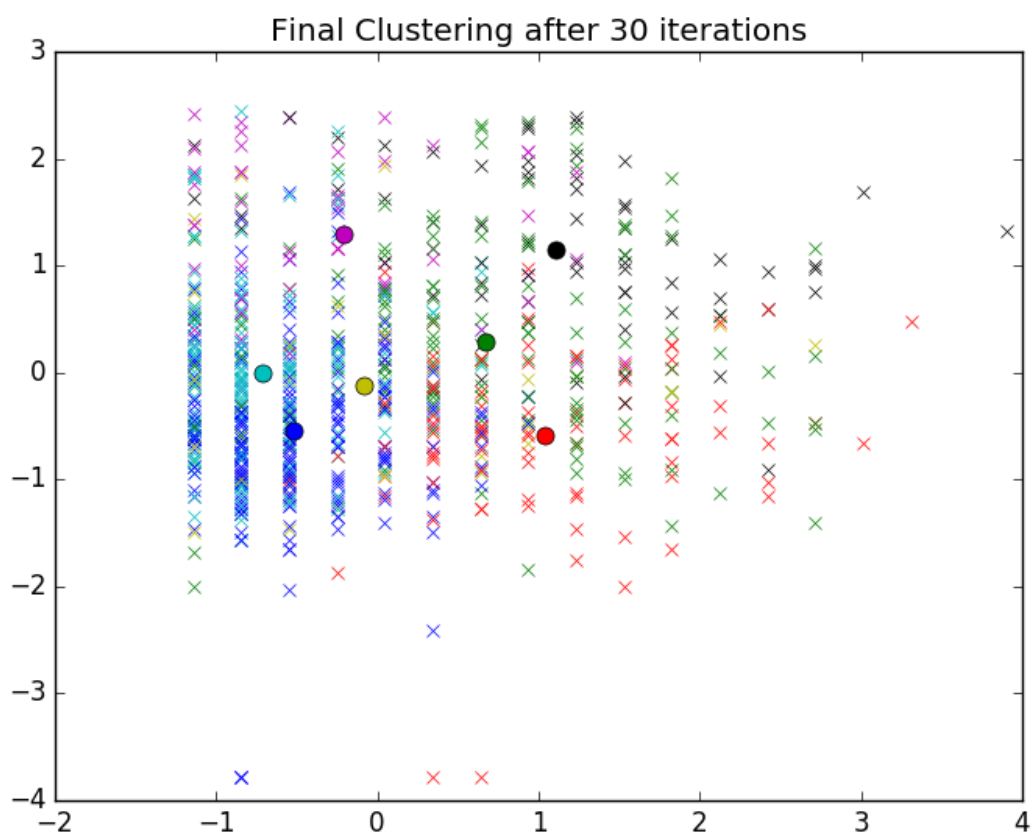


Figure 6. All features, $k=7$, plot features 1, 2