Mark Klobukov

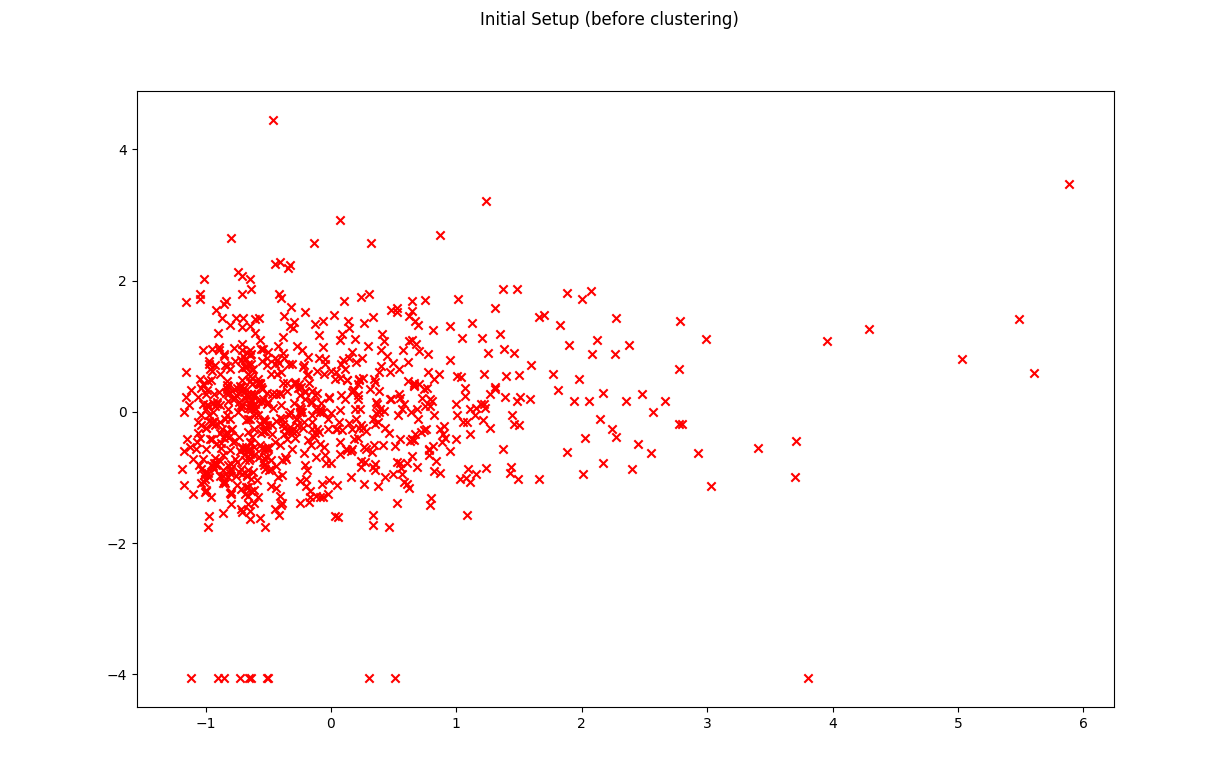
CS 383: Machine Learning

Professor Burlick

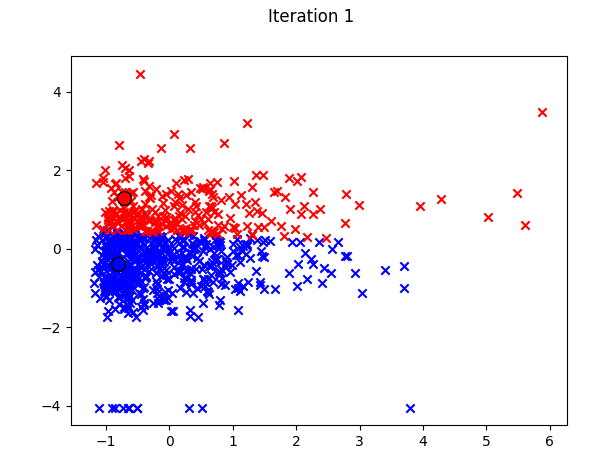
2/1/2018

**Clustering Part I**

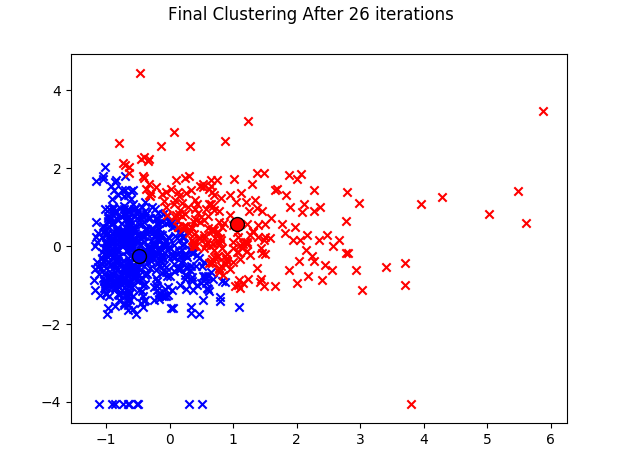
1) Before clustering:

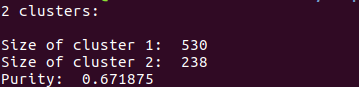


2) Clustering after one iteration of the k-means algorithm:



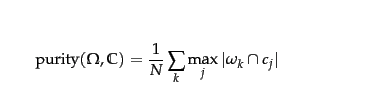
3) Clustering after algorithm termination:





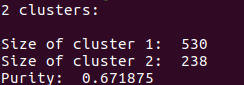
**Purity Calculations:**

Purity was found using the following formula:



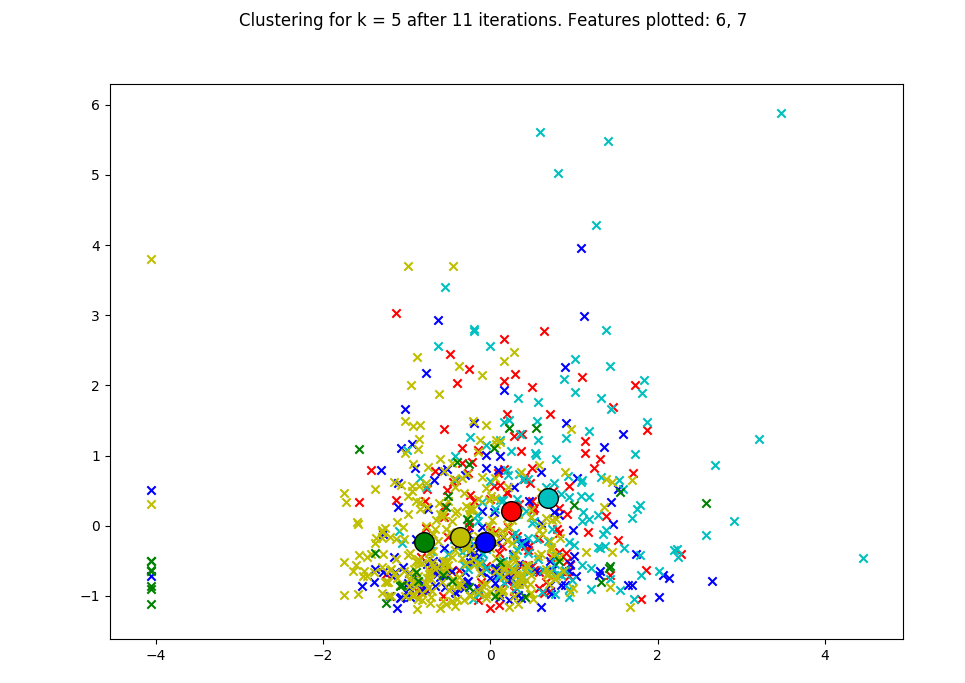
where is cluster k and is class j. In this case, there are two clusters and two classes (the classes correspond to the label -1 or 1 in the first column of the data).

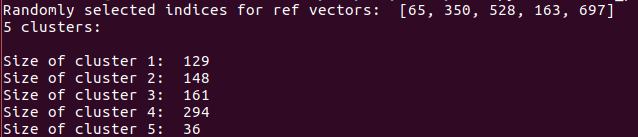
Result of purity calculation is printed to the terminal after running the `kmeans.py` code. Calculated purity value if 0.671875:

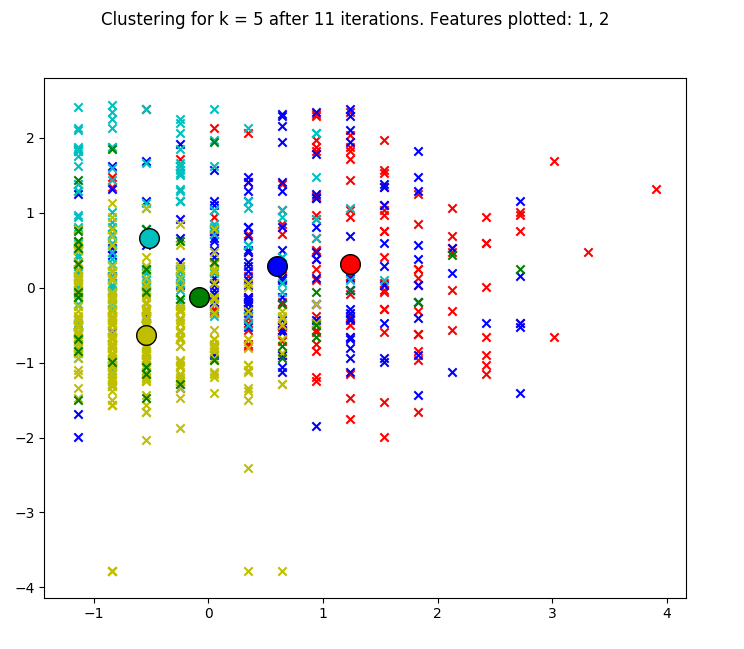


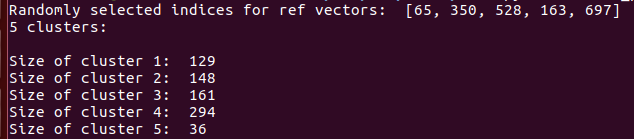
**Clustering Part II**

Example 1: All features used in the algorithm

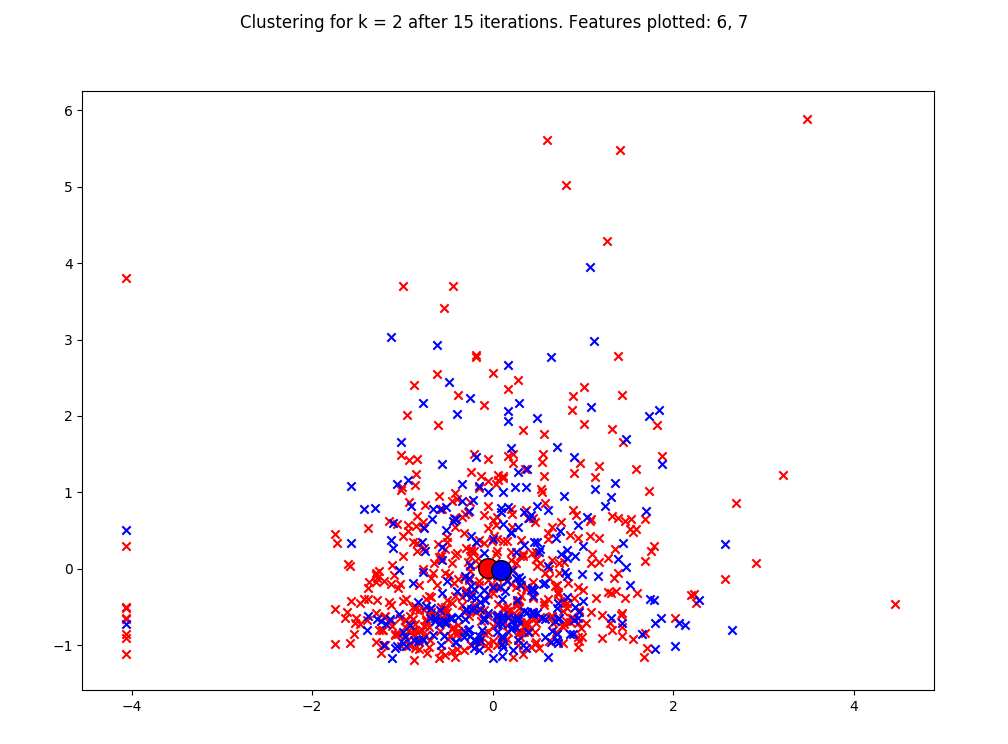


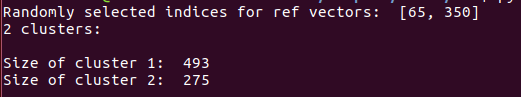


Example 2: All features used in the algorithm 

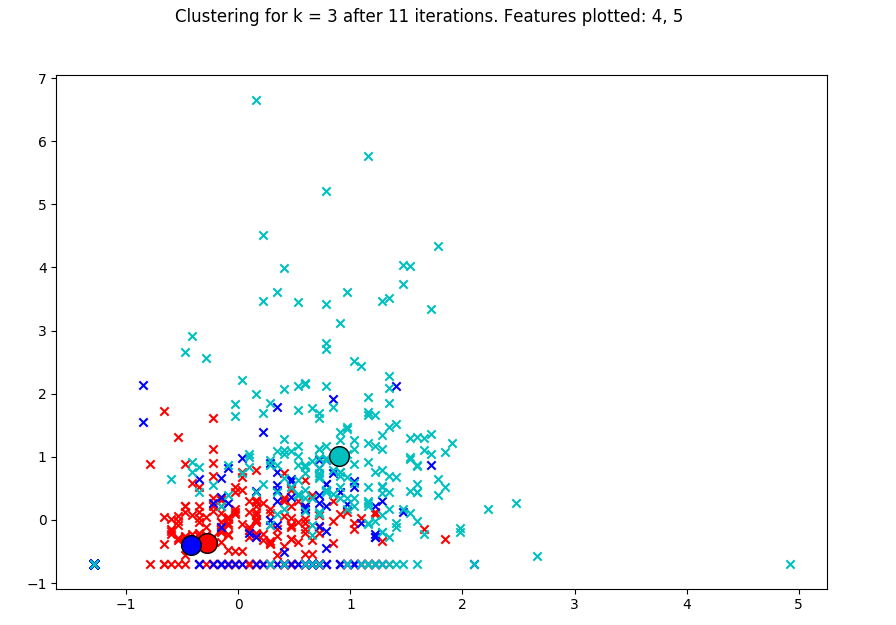


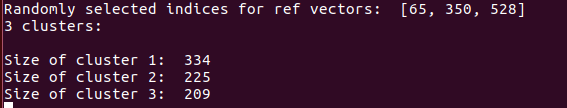
Example 3: All features used in the algorithm





Example 4: All features used in the algorithm





Example 5: All features used in the algorithm

