CS 545

Machine Learning

Programming #3

Submitted by Andrew Huffman

March 13, 2020

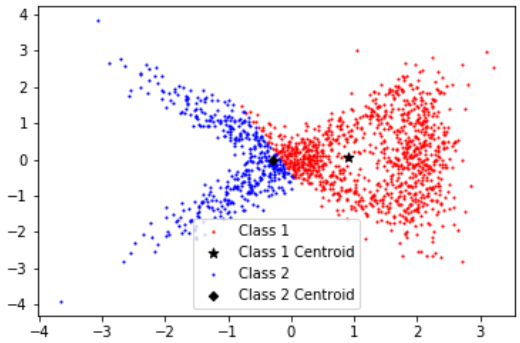
In this assignment I used the K-means and Fuzzy C-means algorithms to partition a 2-D dataset into clusters. The dataset comes from 3 Gaussian distributions and is unlabeled. I ran each testing algorithm for several numbers of clusters (2-5). I ran each combination of clustering algorithm and number of clusters for 10 random initializations. I am reporting the SSW for each of the 10 experiments for each number of clusters, but I am only including plots that show the iterative clustering results for the initialization that yielded the lowest SSW for each number of clusters.

Assignment # 1 K-means:

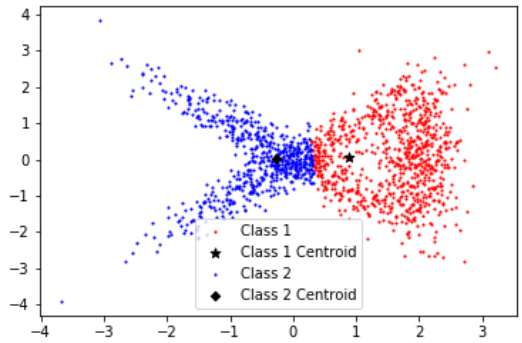
K = 2

r	SSW
0	226.5811
1	363.3516
2	284.4741
3	844.1812
4	226.2393
5	352.2028
6	292.2371
7	286.2566
8	227.14
9	490.3668

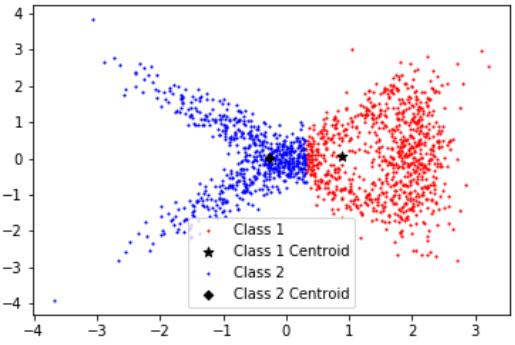
The 5th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 5th initialization of K =2.



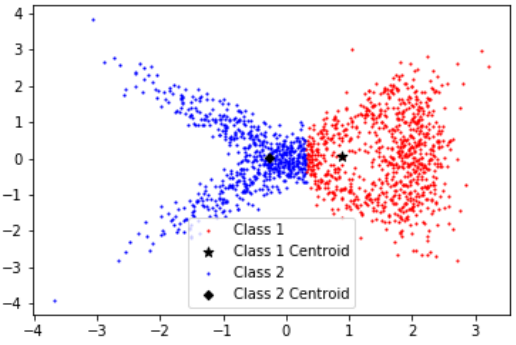
Iteration 1



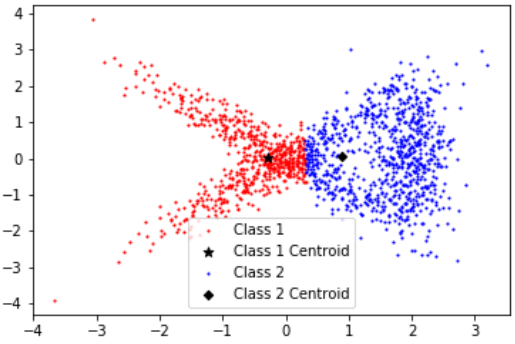
Iteration 2



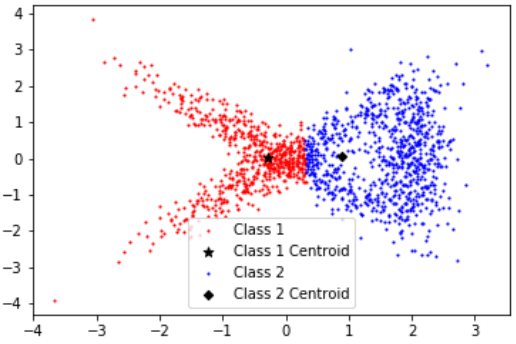
Iteration 3



Iteration 4



Iteration 5

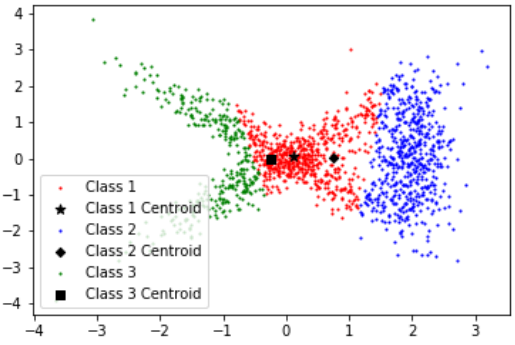


Iteration 6 (Converged)

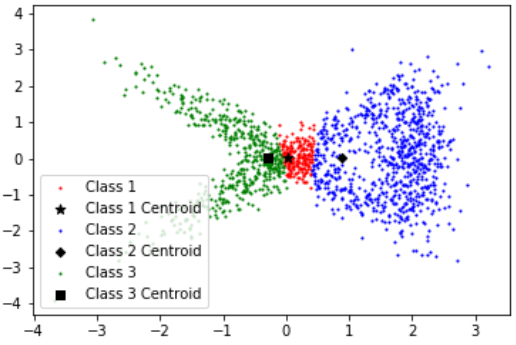
K = 3

r	SSW
0	916.7669
1	603.0147
2	602.4126
3	454.195
4	462.9601
5	500.3263
6	909.6061
7	378.6679
8	863.7188
9	671.1029

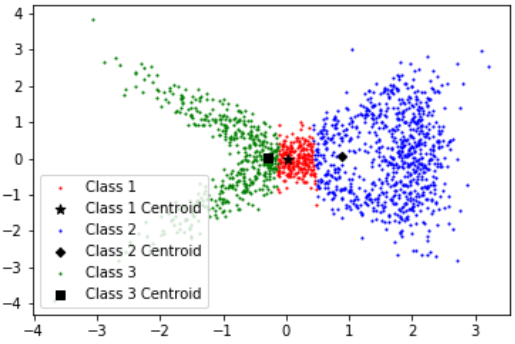
The 8th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 8th initialization of K =3.



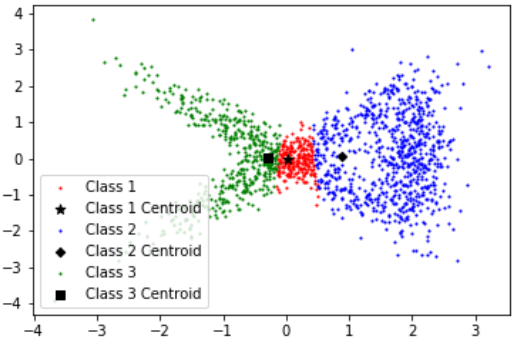
Iteration 1



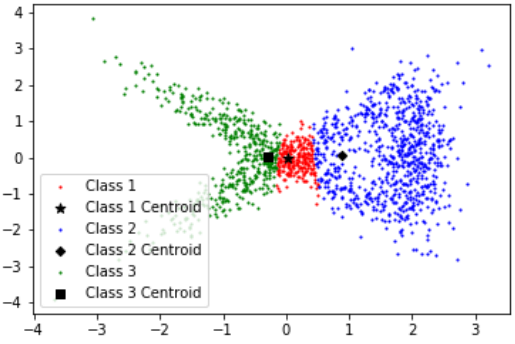
Iteration 2



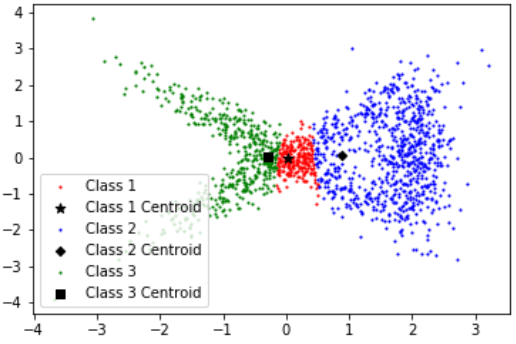
Iteration 3



Iteration 4



Iteration 5



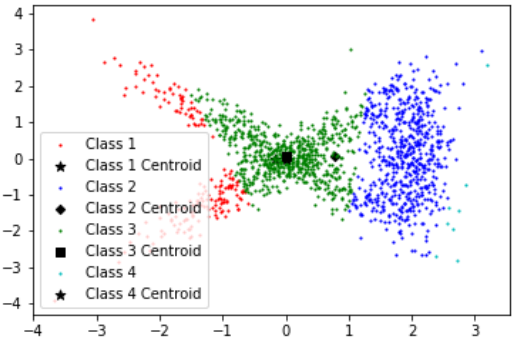
Iteration 6 (Converged)

K = 4

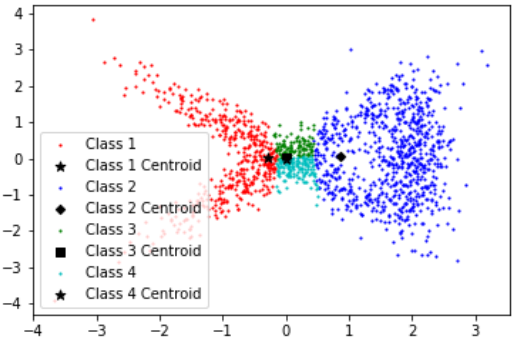
r	SSW

0	653.5796
1	471.5086
2	401.3331
3	474.5989
4	553.3226
5	569.692
6	635.8379
7	636.7041
8	626.3582
9	1084.214

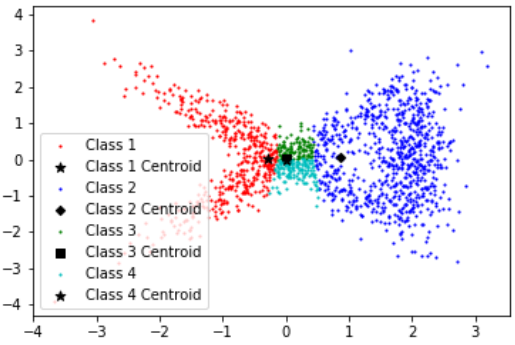
The 3rd initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 3rd initialization of K =4.



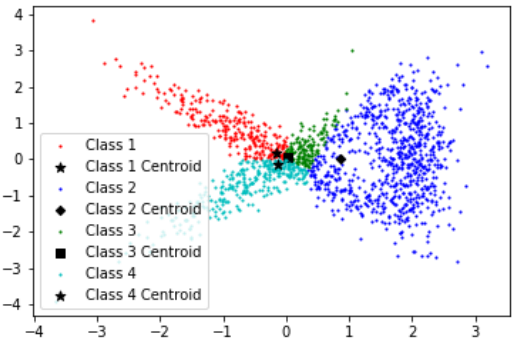
Iteration 1



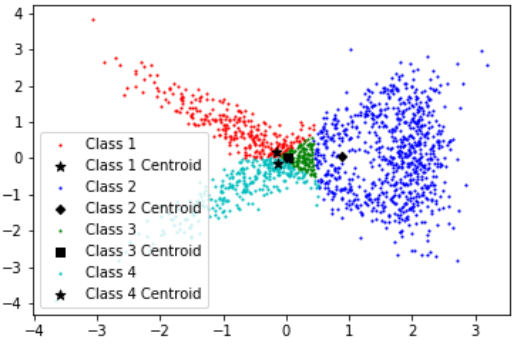
Iteration 3



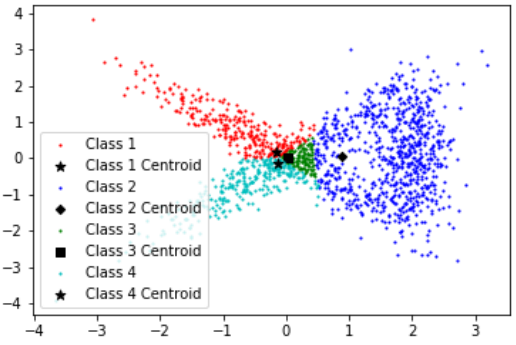
Iteration 5



Iteration 7



Iteration 9

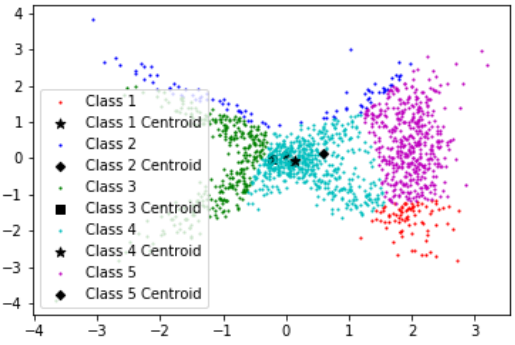


Iteration 11 (Converged)

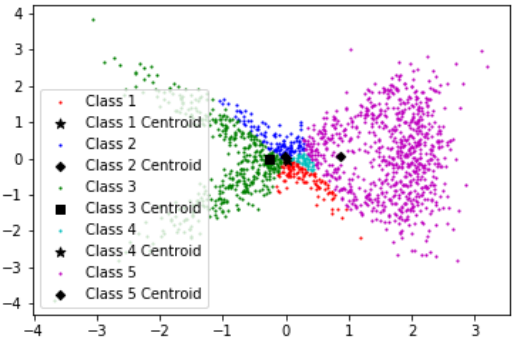
K = 5

r	SSW
0	503.3849
1	2193.864
2	1016.039
3	1202.097
4	588.3366
5	918.1827
6	905.5716
7	1126.473
8	1345.671
9	834.9082

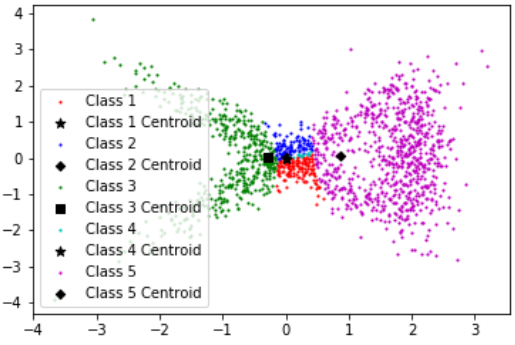
The 0th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 0th initialization of K =5.



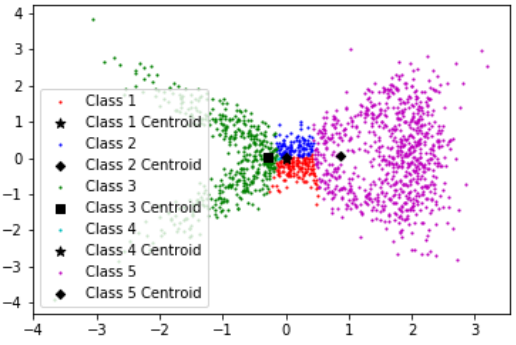
Iteration 1



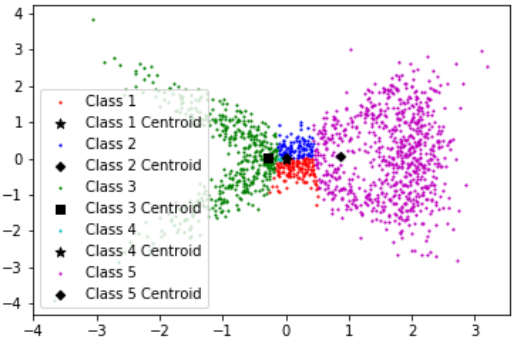
Iteration 2



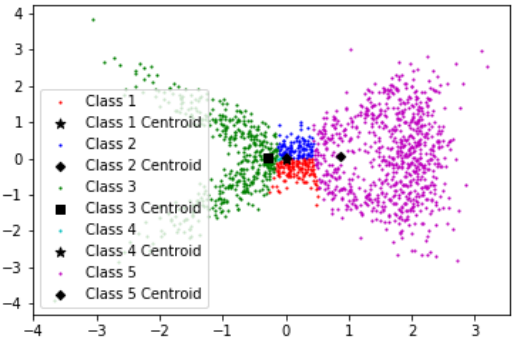
Iteration 3



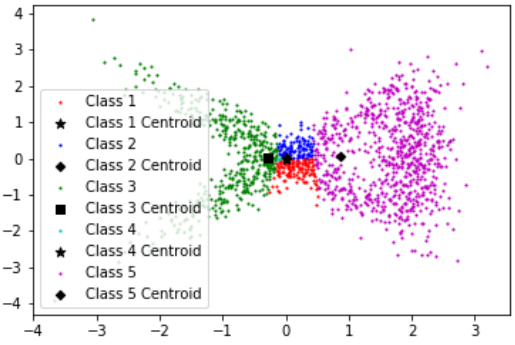
Iteration 4



Iteration 5



Iteration 6



Iteration 7 (Converged)

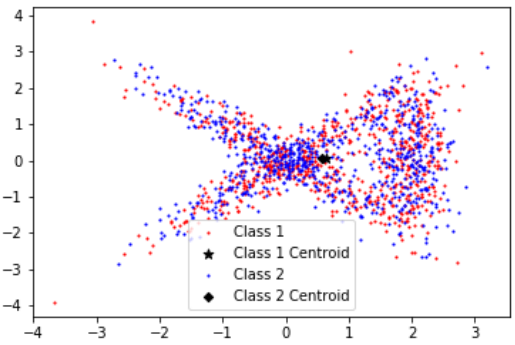
Assignment # 2 Fuzzy C-means:

I used m=2 for these Fuzzy C-means experiments.

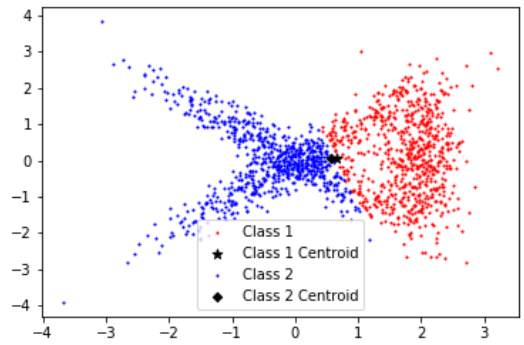
C = 2:

r	SSW
0	702.103
1	923.7319
2	470.8685
3	526.5751
4	701.5073
5	1114.598
6	586.3031
7	526.8215
8	668.863
9	413.992

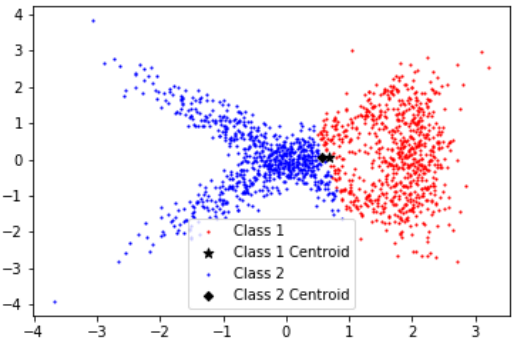
The 9th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 9th initialization of C =2.



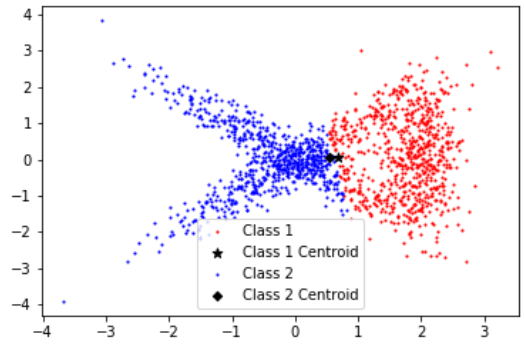
Iteration 1



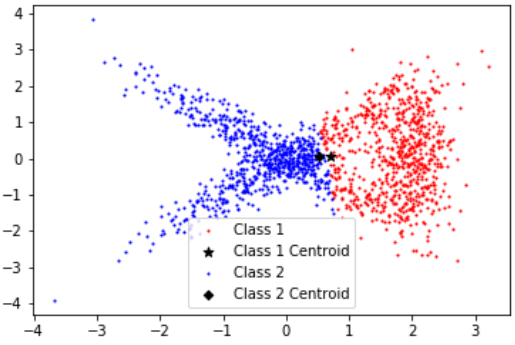
Iteration 2



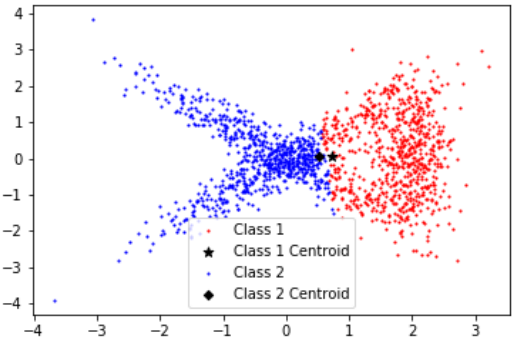
Iteration 3



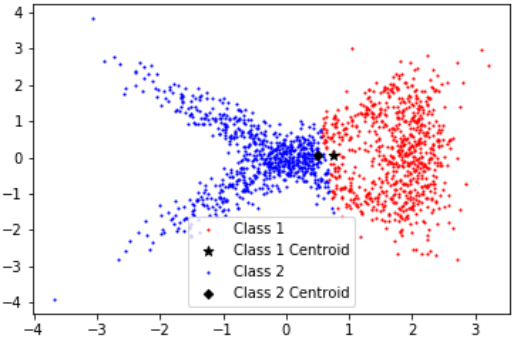
Iteration 4



Iteration 5



Iteration 6

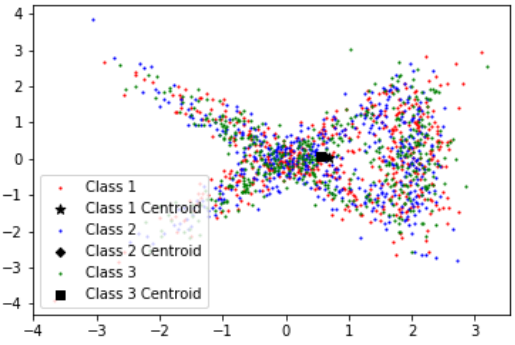


Iteration 7 (Converged)

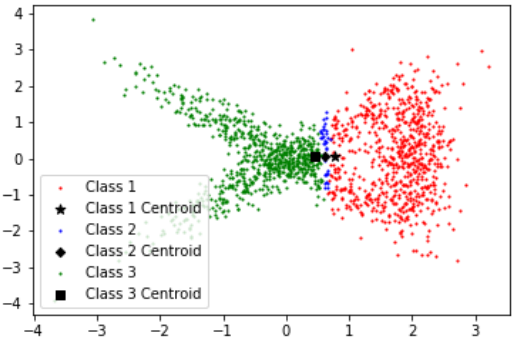
C = 3:

r	SSW
0	1955.679
1	1725.891
2	2119.962
3	1870.437
4	2421.564
5	2266.301
6	1689.499
7	2733.531
8	2282.097
9	2103.441

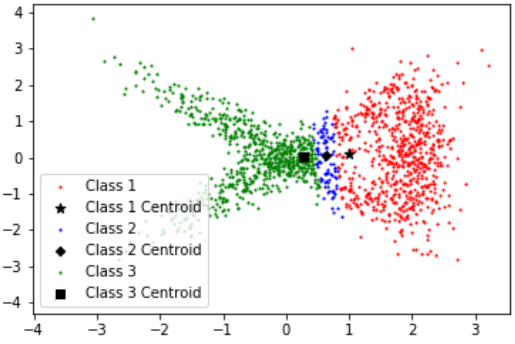
The 6th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 6th initialization of C =3.



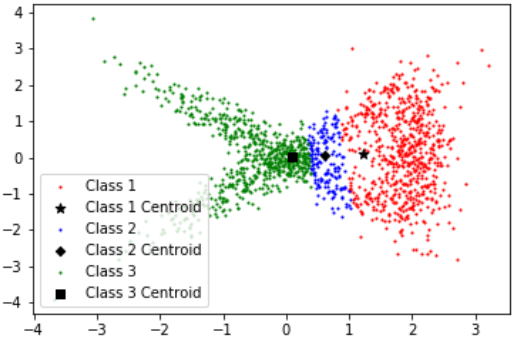
Iteration 1



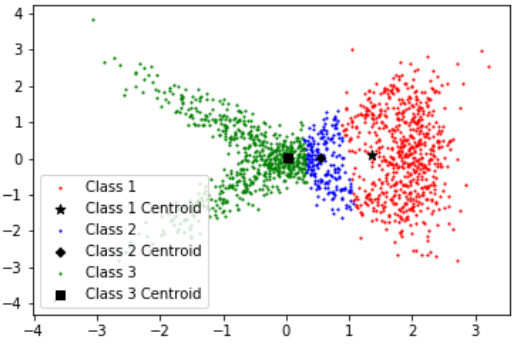
Iteration 5



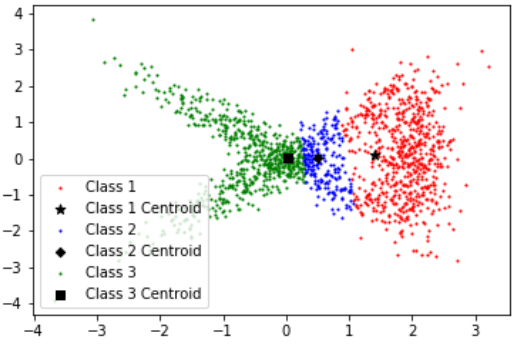
Iteration 10



Iteration 15



Iteration 20

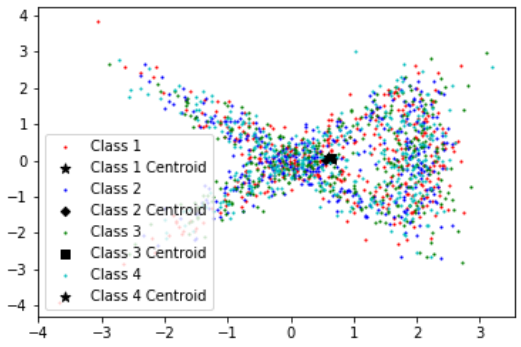


Iteration 26 (Converged)

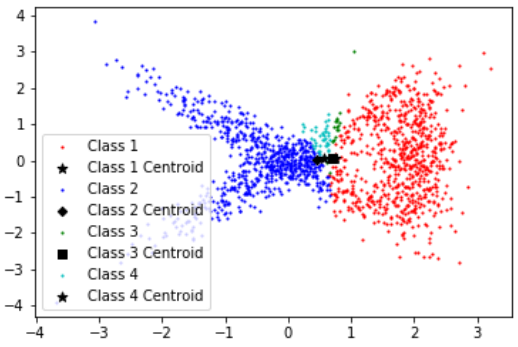
C = 4:

r	SSW
0	1950.76
1	2157.93
2	2796.84
3	2578.53
4	2346.2
5	2291.36
6	2504.16
7	2201.92
8	2020.79
9	2345.98

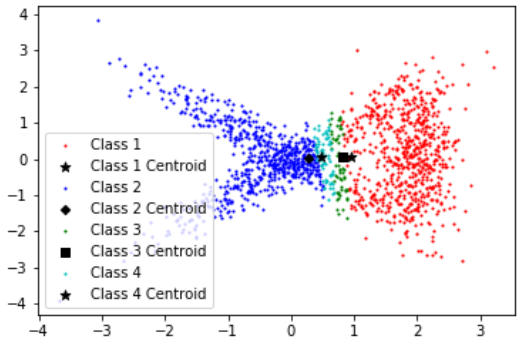
The 0th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 0th initialization of C =4.



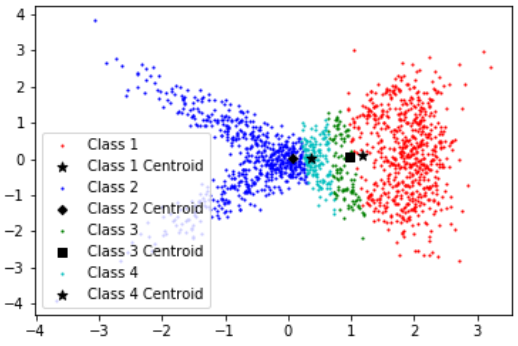
Iteration 1



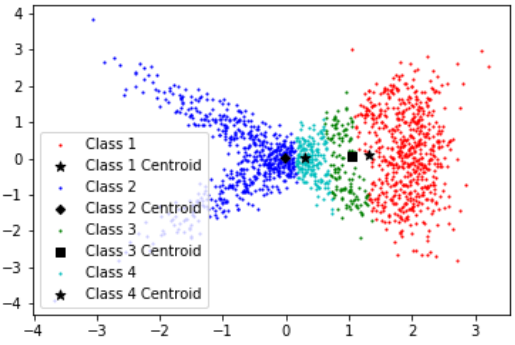
Iteration 5



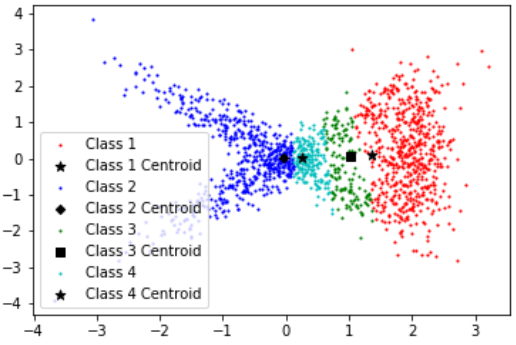
Iteration 10



Iteration 15



Iteration 20

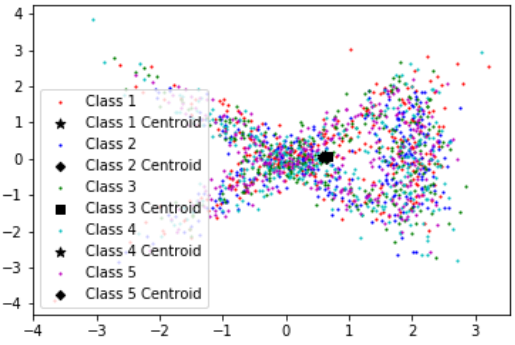


Iteration 27 (Converged)

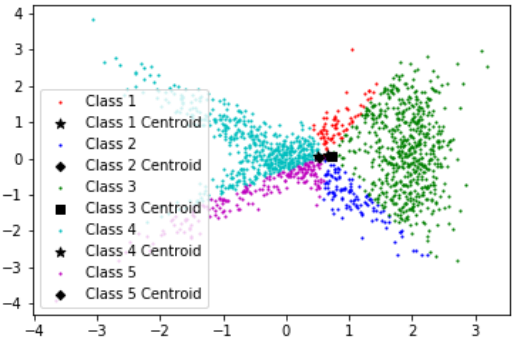
C = 5:

r	SSW
0	3258.08
1	4291.16
2	3674.53
3	4292.36
4	4205.49
5	3178.76
6	3624.45
7	5570.12
8	2816.01
9	4961.33

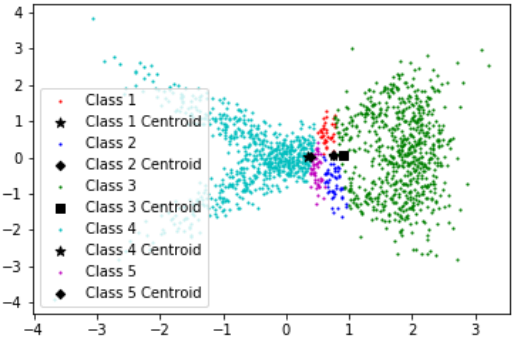
The 8th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 8th initialization of C =5.



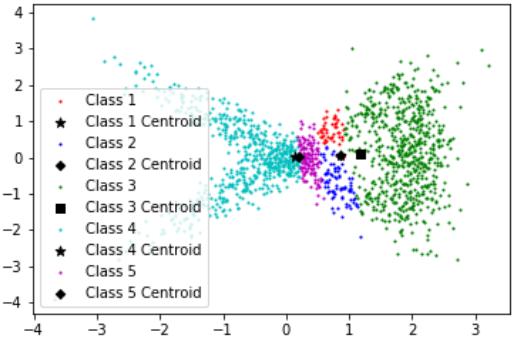
Iteration 1



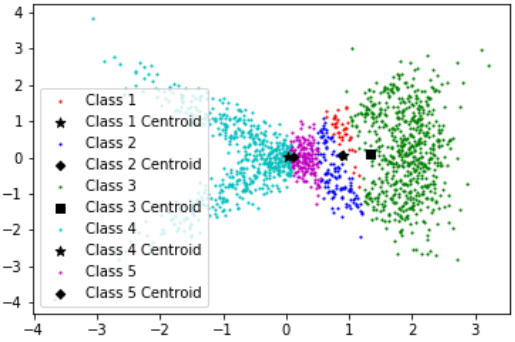
Iteration 5



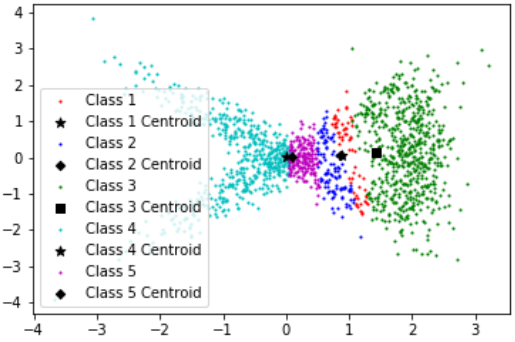
Iteration 10



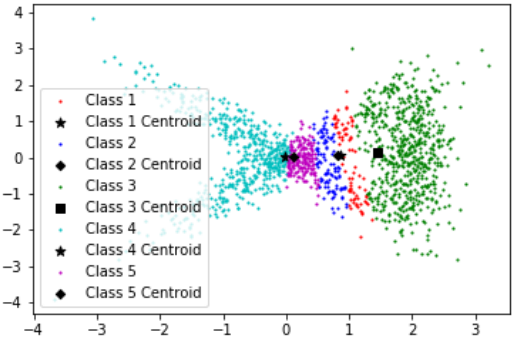
Iteration 15



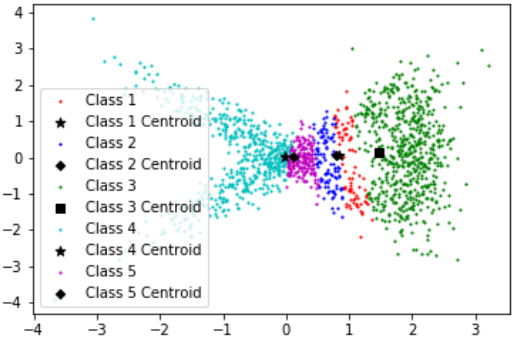
Iteration 20



Iteration 25



Iteration 30



Iteration 34 (Converged)