CS 545

Machine Learning

Programming #3

Submitted by Parth Parashar

PSU ID: - 923928157

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).

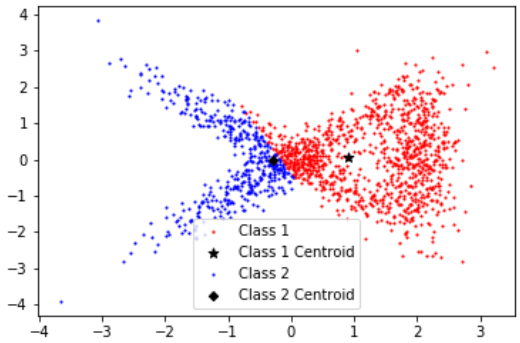
I ran each combination of clustering algorithm and number of clusters for 10 random initializations. I am reporting the SSW for each of the 10 experiments for each number of clusters, but I am only including plots that show the iterative clustering results for the initialization that yielded the lowest SSW for each number of clusters. I will find the correct iterations with the help of SSW(s) yielding the lowest values.

Assignment # 1 K-means:

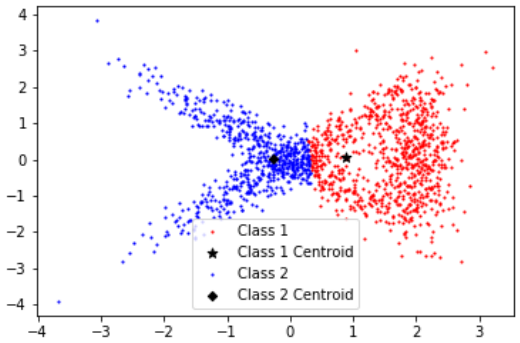
K = 2

r	SSW
0	226.6911
1	364.3716
2	280.4741
3	844.1912
4	226.2393
5	350.1028
6	297.2371
7	281.2566
8	221.14
9	491.3668

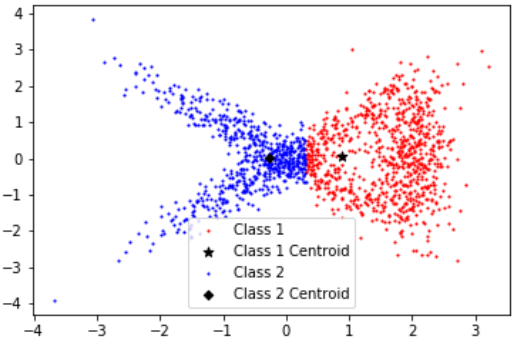
Lowest SSW was obtained in 5th Iteration. Following are plots of several iterations for the 5th initialization of K =2.



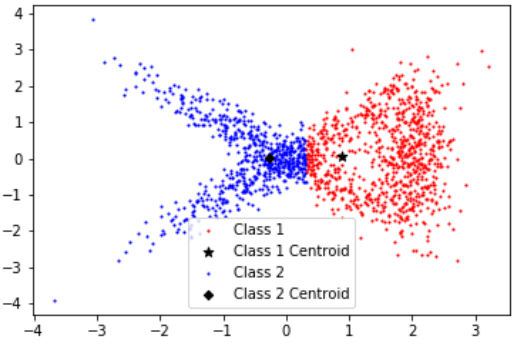
Iteration 1



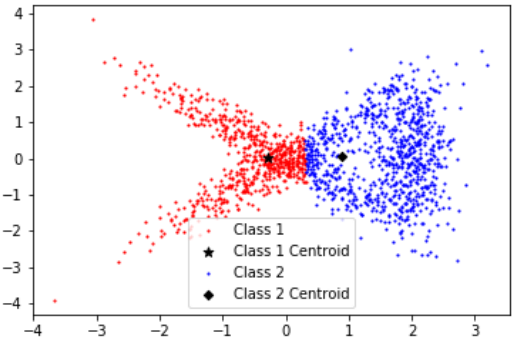
Iteration 2



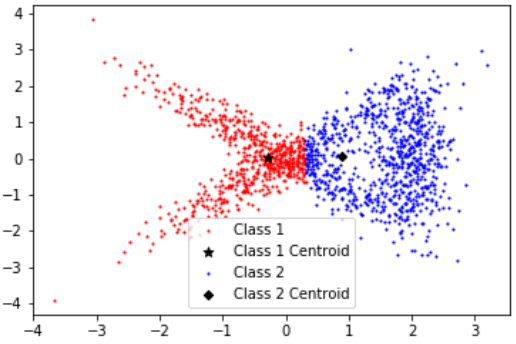
Iteration 3



Iteration 4



Iteration 5

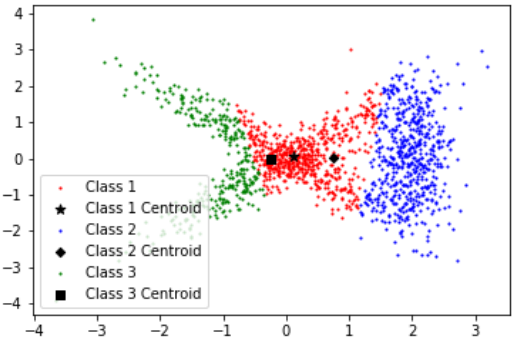


Iteration 6 (Converged)

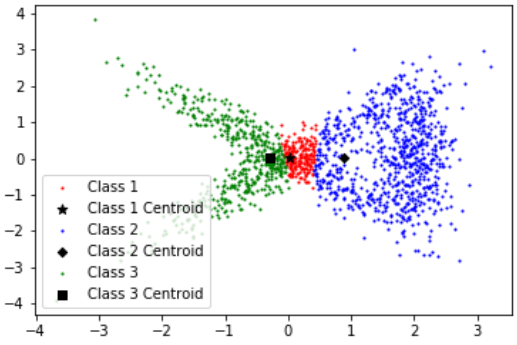
K = 3

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

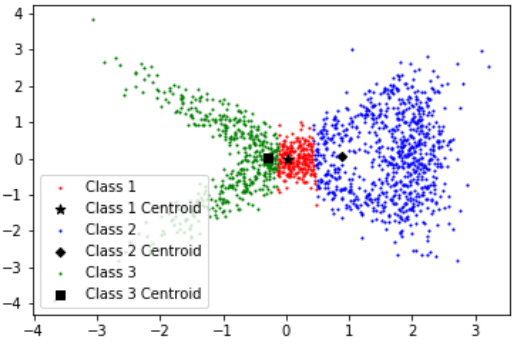
Lowest SSW was obtained on the 8th Iteration. Following are plots of several iterations of the clustering algorithm for the 8th initialization of K =3.



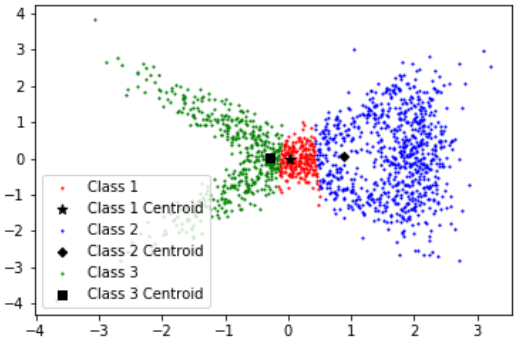
Iteration 1



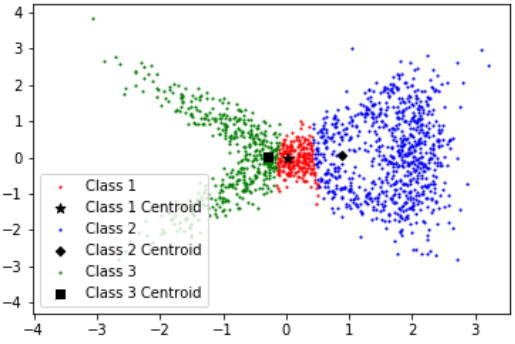
Iteration 2



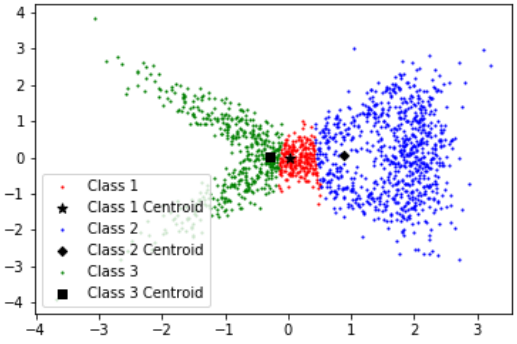
Iteration 3



Iteration 4



Iteration 5



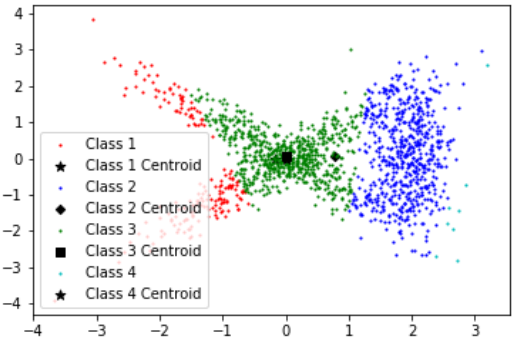
Iteration 6 (Converged)

K = 4

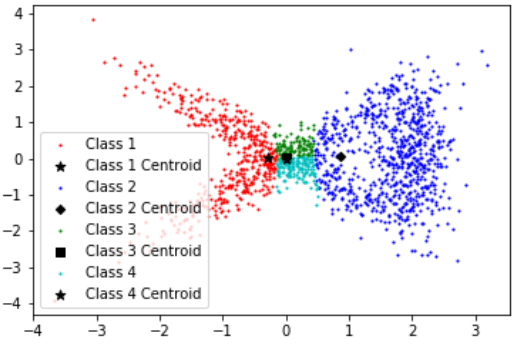
r	SSW

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

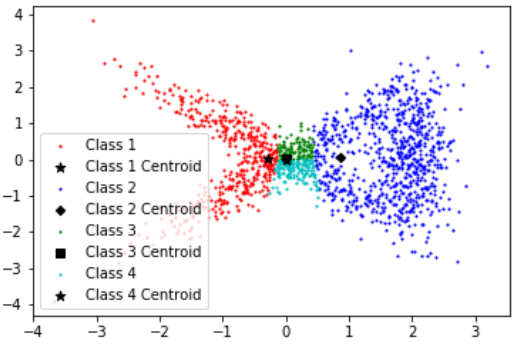
The 3rd initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 3rd initialization of K =4.



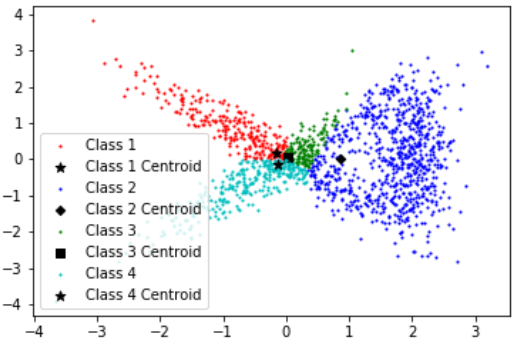
Iteration 1



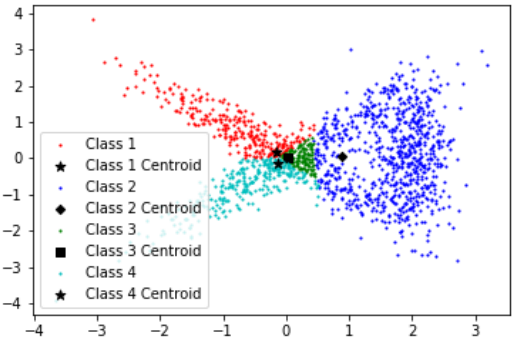
Iteration 3



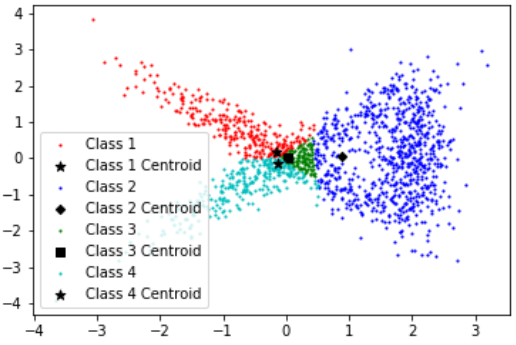
Iteration 5



Iteration 7



Iteration 9

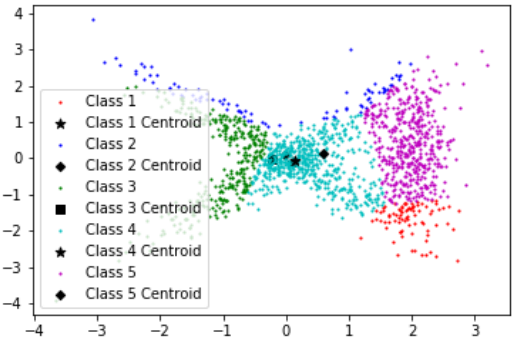


Iteration 11 (Converged)

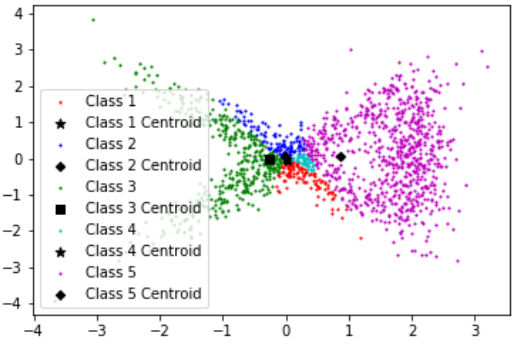
K = 5

r	SSW
0	502.1839
1	2083.854
2	1017.039
3	1201.067
4	589.3133
5	928.1817
6	905.3706
7	1116.463
8	1245.671
9	824.9082

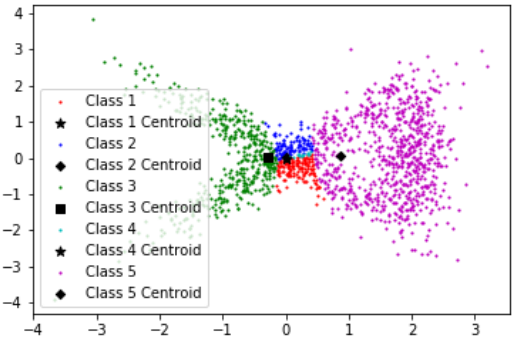
The 0th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 0th initialization of K =5.



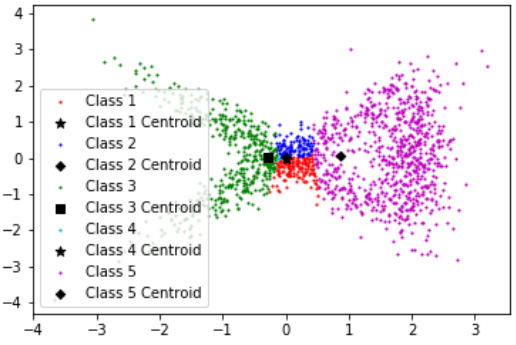
Iteration 1



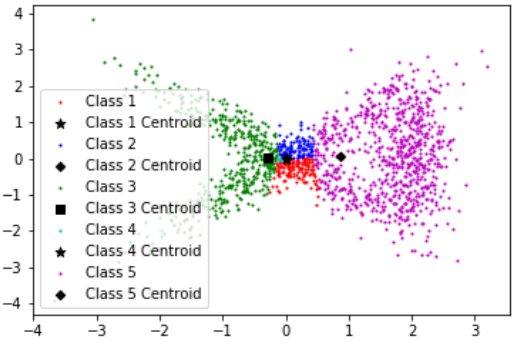
Iteration 2



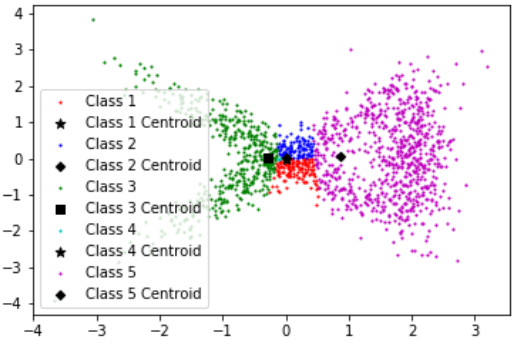
Iteration 3



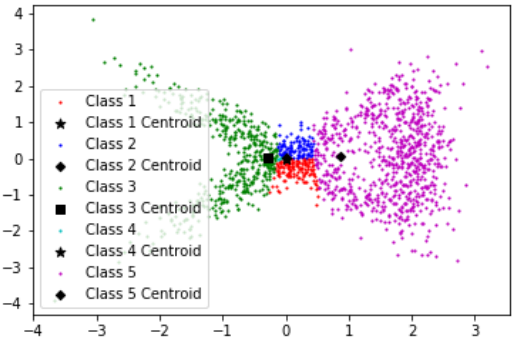
Iteration 4



Iteration 5



Iteration 6



Iteration 7 (Converged)

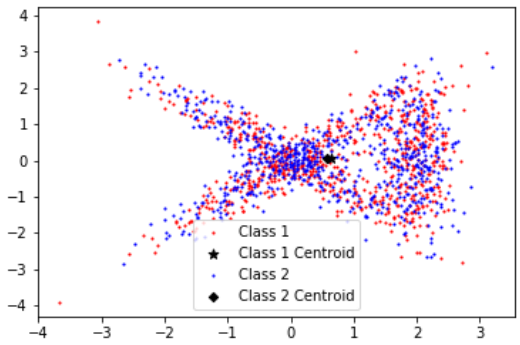
Assignment # 2 Fuzzy C-means:

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

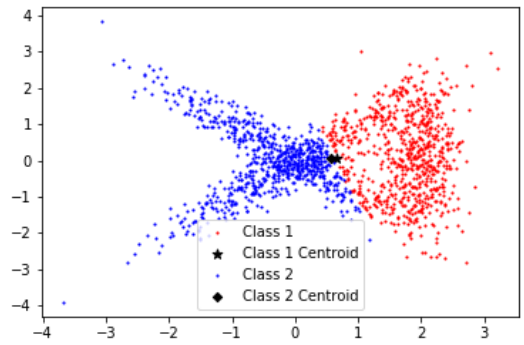
C = 2:

r	SSW
0	712.103
1	924.7319
2	471.8685
3	516.5751
4	701.5073
5	1114.598
6	596.3031
7	516.8215
8	669.776
9	413.992

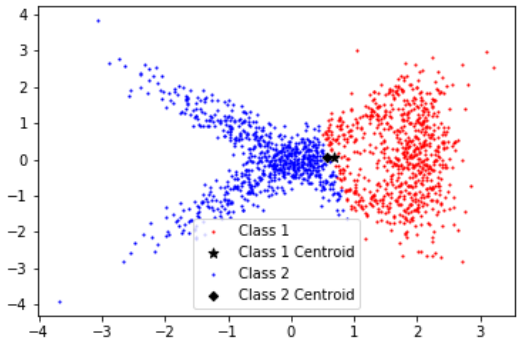
The 9th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 9th initialization of C =2.



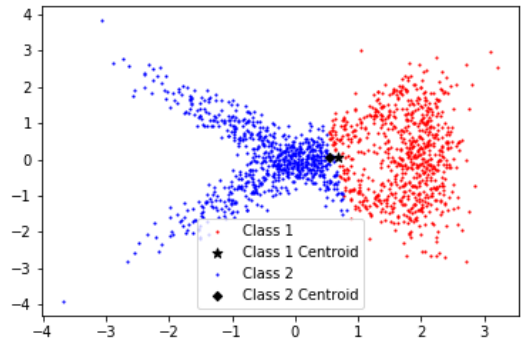
Iteration 1



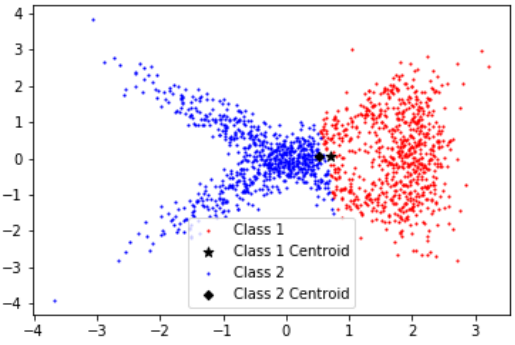
Iteration 2



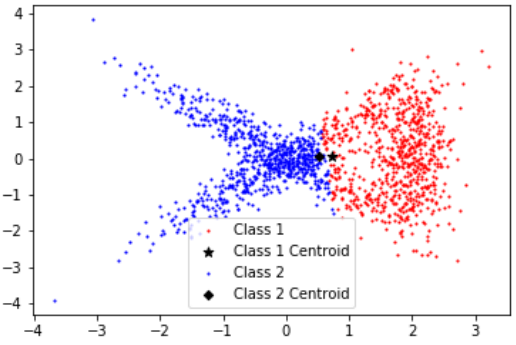
Iteration 3



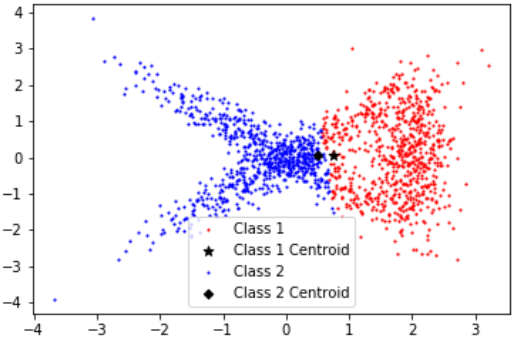
Iteration 4



Iteration 5



Iteration 6

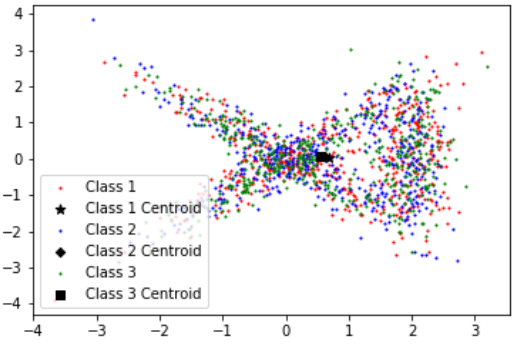


Iteration 7 (Converged)

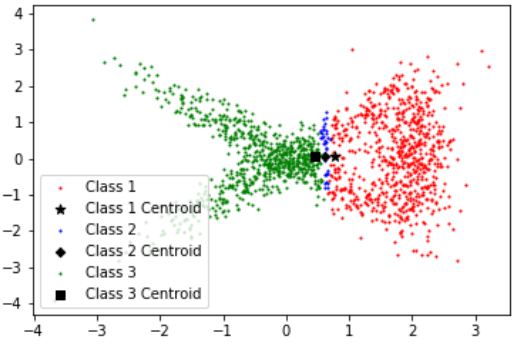
C = 3:

r	SSW
0	1956.579
1	1711.791
2	2109.162
3	1860.417
4	2411.514
5	2269.381
6	1689.499
7	2731.531
8	2262.097
9	2203.441

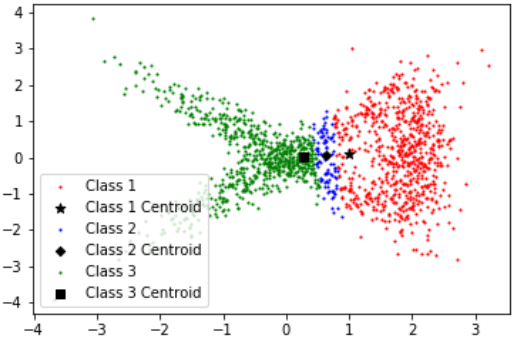
The 6th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 6th initialization of C =3.



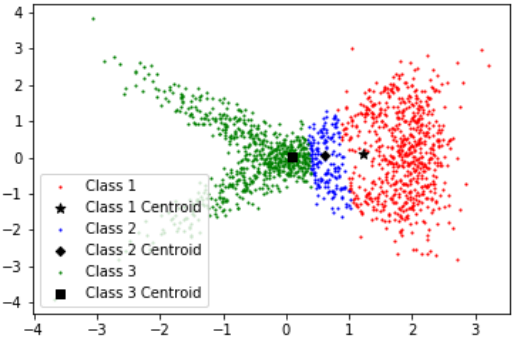
Iteration 1



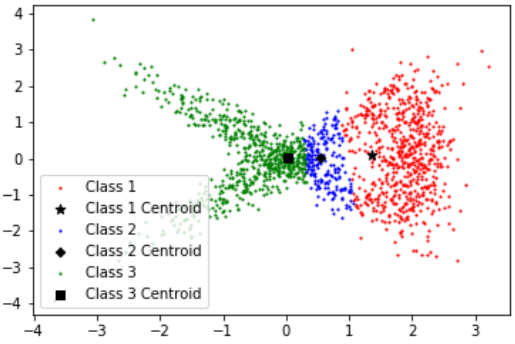
Iteration 5



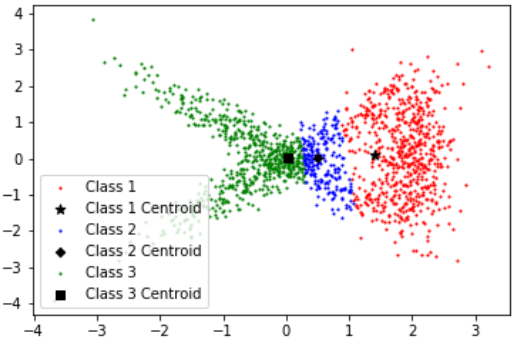
Iteration 10



Iteration 15



Iteration 20

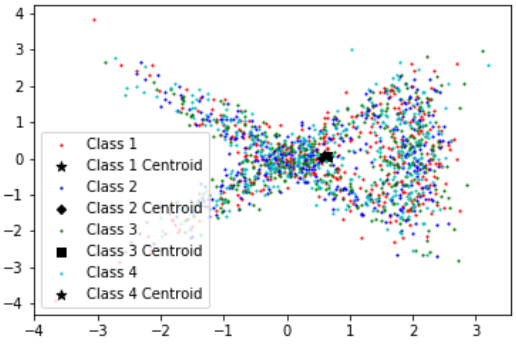


Iteration 26 (Converged)

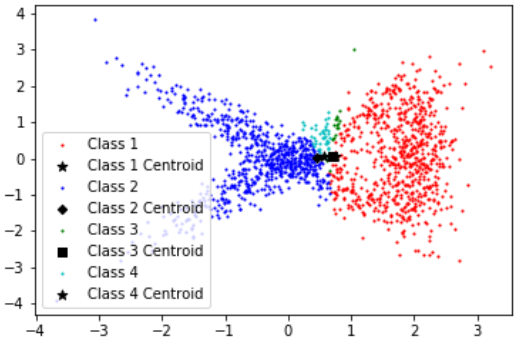
C = 4:

r	SSW
0	1950.76
1	2147.83
2	2797.94
3	2579.63
4	2347.2
5	2191.36
6	2501.16
7	2202.92
8	2021.79
9	2348.98

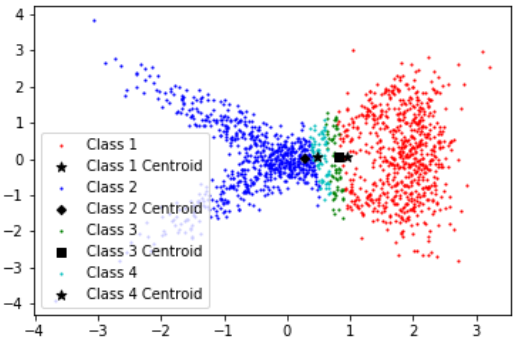
The 0th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 0th initialization of C =4.



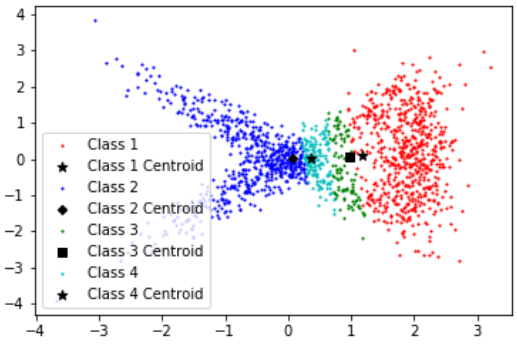
Iteration 1



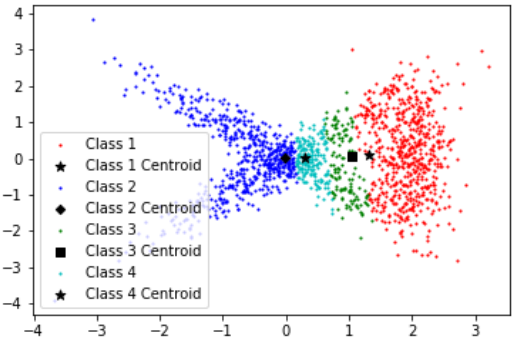
Iteration 5



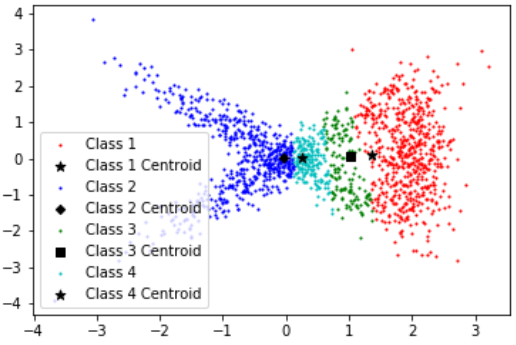
Iteration 10



Iteration 15



Iteration 20

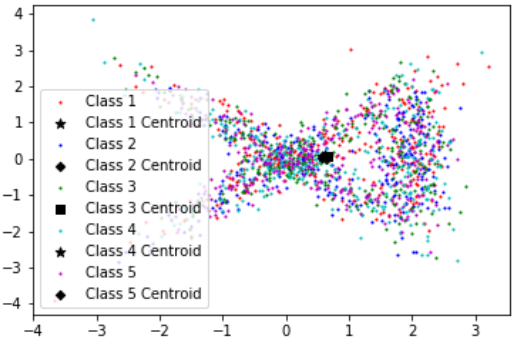


Iteration 27 (Converged)

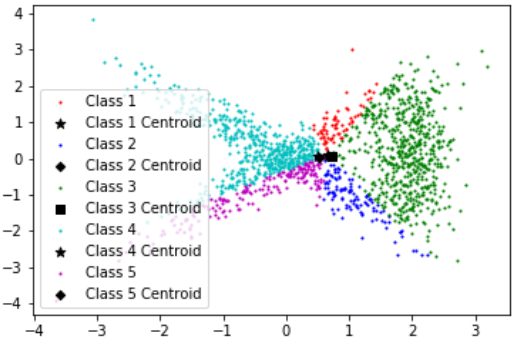
C = 5:

r	SSW
0	3259.08
1	4190.16
2	3663.73
3	4261.36
4	4273.59
5	3778.16
6	3621.85
7	5210.11
8	2816.01
9	4861.63

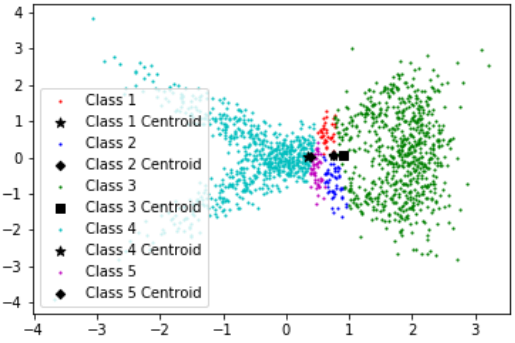
The 8th initialization gave the lowest SSW. Following are plots of several iterations of the clustering algorithm for the 8th initialization of C =5.



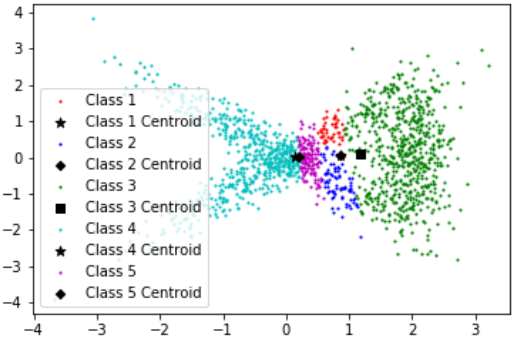
Iteration 1



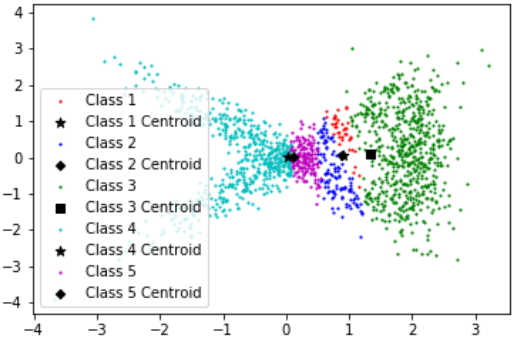
Iteration 5



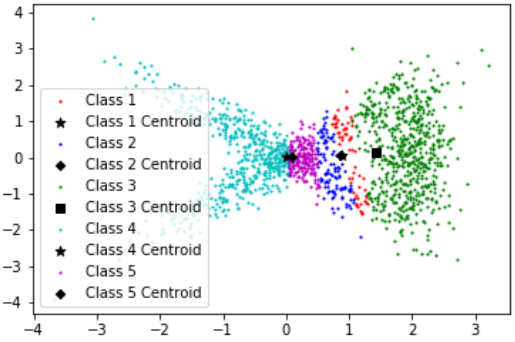
Iteration 10



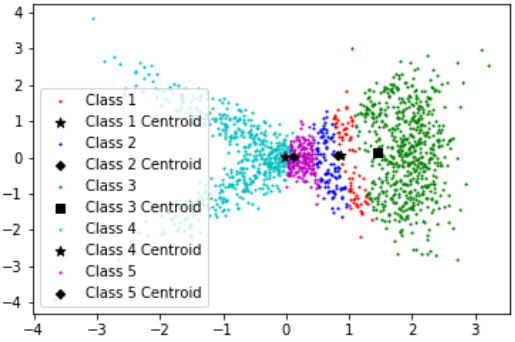
Iteration 15



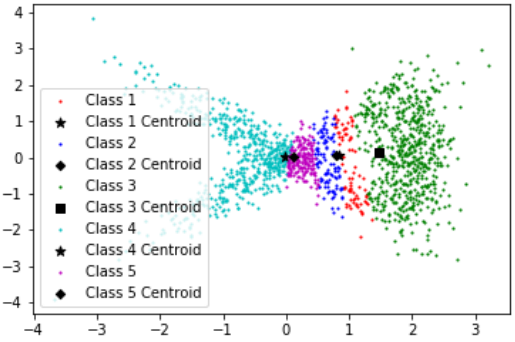
Iteration 20



Iteration 25



Iteration 30



Iteration 34 (Converged)