Machine Learning Homework 6

k-means clustering, kernel k-means, spectral clustering, DBSCAN

Due day 23:59 PM 3rd June

- **Homework Objective:** Use whatever your favorite language to code out k-means clustering, kernel k-means, spectral clustering and DBSCAN. You are allowed to use NumPy and it can help you to solve the eigenvalue problem. Using the matplotlib is legal. **But you can't use scikit-learn and SciPy in this homework.**
  (kernel k-means and spectral clustering both based on RBF kernels)

- **Dataset:** 2 datasets with points on 2d space, **circle.txt** and **moon.txt**

- **Requirements:**
  ‣ **20 points** : You need to make videos or GIF images to show the clustering procedure (visualize the cluster assignments of data points in each iteration, colorize each cluster with different colors) of your k-means, kernel k-means, spectral clustering and DBSCAN program.

  ‣ **15 points** : In addition to cluster data into 2 clusters, try more clusters (e.g. 3 or 4) and show your results.

  ‣ **15 points** : For the initialization of k-means clustering used in k-means, kernel k-means and spectral clustering, try different ways and show corresponding results, e.g. k-means++.

  ‣ **15 points** : For spectral clustering, you can see if data points within the same cluster do have the same coordinates in the eigenspace of graph Laplacian. You should plot the result and discuss it in the report.

  ‣ **35 points** : Submit a report with showing your code and give detailed explanations. You should explain every thing you have done in this homework and show all your results in the report. The report should be written in English.

- **Turn in:**
  1. Report (.pdf)
  2. Source code
  You should zip source code and report in one file and name it like ML_HW6_yourstudentID_name.zip, e.g. ML_HW6_0756504_雷承勳.zip.

  P.S If the zip file name have format error, it will be penalty (-10). Please submit your homework before deadline, late submission is not allowed.