CMPE 255-02, Spring 2024

Assignment #2

Release on March 6th, 2023
Due 11:59pm on Sunday, March 17th, 2023

Notes

This assignment should be submitted in Canvas as a format of ipython notebook (assignment2_yourFirstName_LastnName.ipynb).

No late assignments will be accepted.

You may collaborate on homework but must write independent code/solutions. Copying and other forms of cheating will not be tolerated ~~~~~~ (imal penalty) or a failing grade for the course. Your w ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ ess, and clarity, not just the answer. Thus, correct answe ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ eive very little credit.

NOTE: Please do not use any p ~~~~~~ Pandas, and Matplotlib.

Please download cluster_dat ~~~~~~~~~~~~

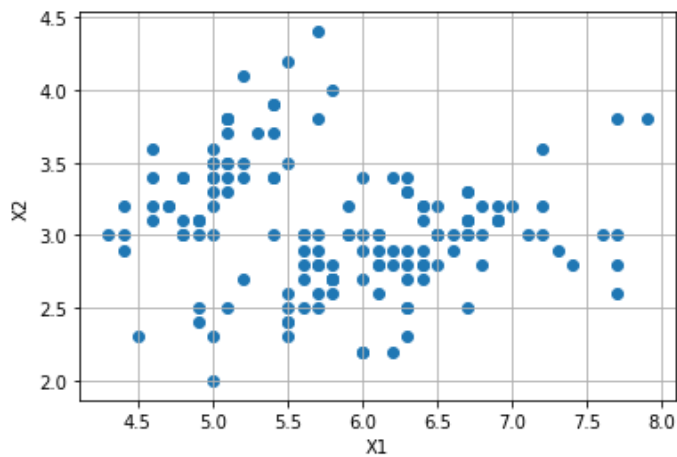| X1	X2	X3	
6.7	3	5	
6.3	2.9	5.6	
5.6	3	4.5	
7.6	3	6.6	
6	3.4	4.5	
6.4	3.2	5.3	
7.7	2.8	6.7	
4.8	3	1.4	
5	3	1.6	
5	3.4	1.6	0.4

1. (4 pts) Implanting K-means clustering algorithm

K-means algorithm is a method to automatically cluster similar data examples together. K-means is an iterative procedure that starts by guessing the initial centroids, and then refines this guess by repeatedly assigning examples to their closest centroids and then recomputing the centroids based on the assignments until converge.

Let's assume K=3. Please implement K-means clustering algorithm from scratch.
Please plot the location of k centroids like the below figure for the first 5 steps including the initial setting of the centroids.
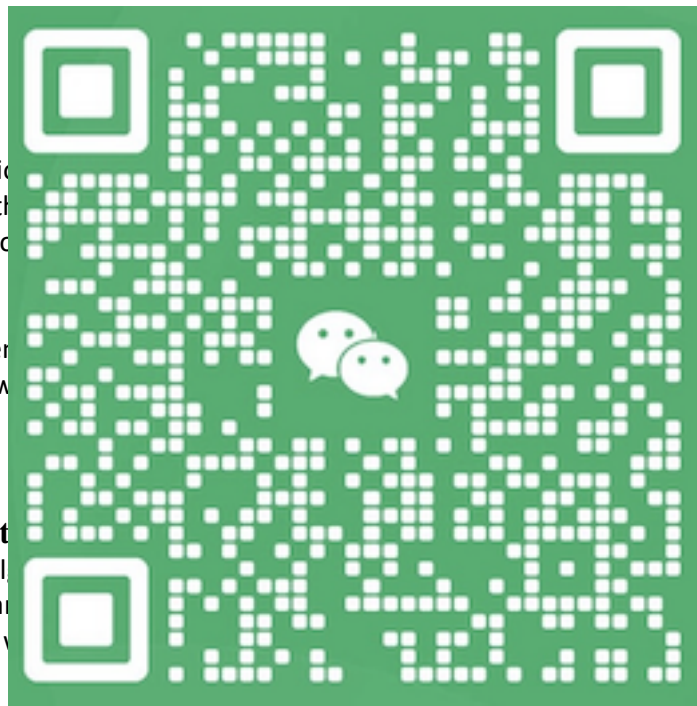

2. (2 pts) Run a few trials

Note that the converged solutio[n] ... [ng of the centroids. Therefore,] in practice the K-means algorith[m] ... [t]ations. One way to choose between these different solutio[ns] ... [th]e with the lowest objective function value.

Please run 5 times with differe[nt] ... [cal]culate the objective function values. Then, choose the trial w[ith ...]

3. (4 pts) K-mean++ algorit[hm]

Please implement K-mean++ al[gorithm ...]
Please plot each step of K-mea[n ...]
Please compare current result [with ...]