

Homework **xxx**

CSE 411

K-means Clustering and GUI in Matlab

In this homework, you will write a Matlab application that performs K-means clustering. As we learned in class, K-means clustering algorithm partitions n data points into k clusters ($n \gg k$) with each data point belongs to the cluster with the nearest mean (or centroid). In this homework, you will practice data generation, data clustering, and display of clustering results in a Matlab GUI created using GUIDE.

Specifically, the steps you should follow to complete the homework are as follows.

1. Data generation. You have several choices:
 - Generate k sets of data from k normal distributions of different mean and standard deviation. (Matlab function hint: randn)
 - Use Matlab function "ginput" to allow the user to directly input/click data points in a 2D axes.
 - (Optional) add an Edit box or a Popup Menu or a Slider to allow the user to enter the value for k .
 - (Optional) have edit boxes or sliders to enter the values of mean and standard deviation for each set of generated data.

After the data are generated, display all data points in a figure/axes in your GUI.

2. K-means clustering. Add a pushbutton to your GUI. When the user clicks on the pushbutton, its callback function should perform K-means clustering algorithm, save clustering result in the Matlab struct "handles", and also display the clustering result in the same or a separate figure/axes in your GUI. In the display of result, data points in each cluster should be displayed with a different color (one color for one cluster).

Beyond the above basic requirements, you are encouraged to add more UIcontrols, alternative data generation or data clustering mechanisms to your program. For instance, one extension of your program could be the animation of the clustering process, by displaying the intermediate results after each iteration in the K-means clustering algorithm. (Hint: to avoid the results being refreshed too fast to see, you can use Matlab function "pause(m)" to pause *m* seconds before continuing execution and display of the result of another iteration.)

Notes:

Matlab provides an implementation of K-means clustering (kmeans). Please do not use this function for this assignment. You can use the kmeans function as reference however, or compare the result of your implementation with the result of the Matlab implementation.

References:

<http://www.mathworks.com/help/matlab/>

<http://www.mathworks.com/help/matlab/gui-building-basics.html>

http://www.mathworks.com/help/matlab/creating_plots/using-high-level-plotting-functions.html

Questions?

Email Prof. Huang at xih206@lehigh.edu, or come to her office hours on Tuesday 3:00-5:00pm in Packard Lab 218.