Name: Nakshatrala Hari Hara Kumar
Student ID: 1001102740

Python Version:
Python 3.6

Libraries Version:
Numpy – 1.14.3
Matplotlib – 2.2.2

Code Explanation:
The code can be explained in following steps.

Data reading:
First, we need to load the data set into two numpy arrays called x_matrix and y_matrix

Initialize Centroids:
Then as the value of k mentioned by the user, we need to generate the initial points of centroids. Here I took the min and max value of each dimension as the range of random number generated.

Calculate the Euclidian distance:
Then, we need to calculate the Euclidian distance between each data point to each cluster and assign each point to nearest cluster.

Change the cluster to new location:
Once we have the points assigned to each cluster then we need to move the centroid to the center of all the points under that cluster.

Repeat:
Repeat this whole process for 300 times to get the best possible solution

Calculation of accuracy:

Once we have all the points to centroid, science the data we have is a labelled data we can calculate the accuracy by seeing how many points are wrongly classified.

How to run the code:
1. Place the text file with the data in the same location

2. Run the file Main.py with k value as the parameter

Python Main.py <k-value>