Question 2.

I have executed the program on the iris.txt database for 10 times and calculated the purity score every time.

For this, I made the function getPurityScore() in my program to get the purity score.

I executed the program for 10 times with different random centroids and output the purity score, and finally got the best purity score 0.92

But the purity scores and maximum purity were different every time.

The experiment shows that the result for clustering is affected by the selection of centroids, so it is very important to select the proper centroids.

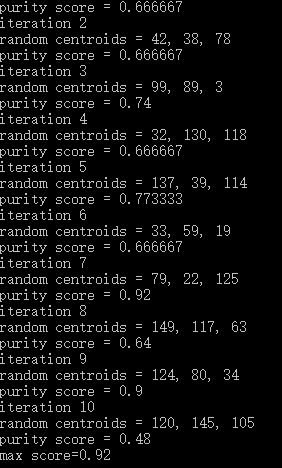


Figure 1. Purity scores and maximum purity score from the execution of 10 times with different random centroids.