**K-Means Clustering using SSTree Technique(2 levels)**

**Implementations Detail:**

|  |  |
| --- | --- |
| 1 | Generated Random n data points. |
|  | **Level 1 clustering** |
| 2 | Computer k cube root of number of data points. |
| 3 | Selected K random clusters among data points as the cluster centroids using farthest distance method. |
| 4 | Assigned data points to the nearest cluster. |
| 5 | Calculated centroids for each cluster based on data point assignment. |
| 6 | Repeated step 3 and 4 until cluster centroid does not change. |
| 7 | Sort the data based on cluster assignment of level 1. |
|  | **Level 2 clustering** |
| 8 | Computer k1 cube root of number of data point in Kth cluster. |
| 9 | Selected k1 random clusters among data points as the cluster |
| 10 | Assigned data points to the nearest cluster. |
| 11 | Calculated centroids for each cluster based on data point assignment. |
| 12 | Repeated step 8 and 9 until cluster centroid does not change. |
| 13 | Sort the data based on cluster assignment of level2. |
| 14 | Randomly generate Search data point. |
| 15 | Search level 1 cluster K, if search data point fall’s within cluster K’s radius. Search sub cluster’s of Kth cluster if data point fall’s within cluster K’s sub clusters radius: provided K’s sub cluster size is greater than Zero and perform exhaustive search in that sub cluster. |

**Instructions to run the code:**

**How to compile**

gcc -c "cs5331-HW4-(Arun kumar Jegarkal)-(Jaichandra Sesetty)-(Shivendra Satish Dubeer).c"

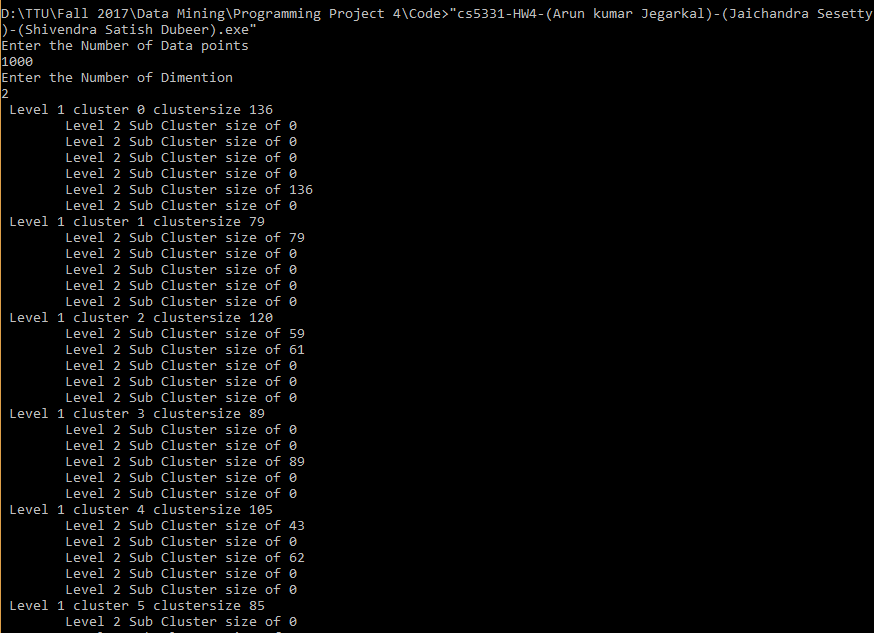
gcc "cs5331-HW4-(Arun kumar Jegarkal)-(Jaichandra Sesetty)-(Shivendra Satish Dubeer).o" -o "cs5331-HW4-(Arun kumar Jegarkal)-(Jaichandra Sesetty)-(Shivendra Satish Dubeer).exe"

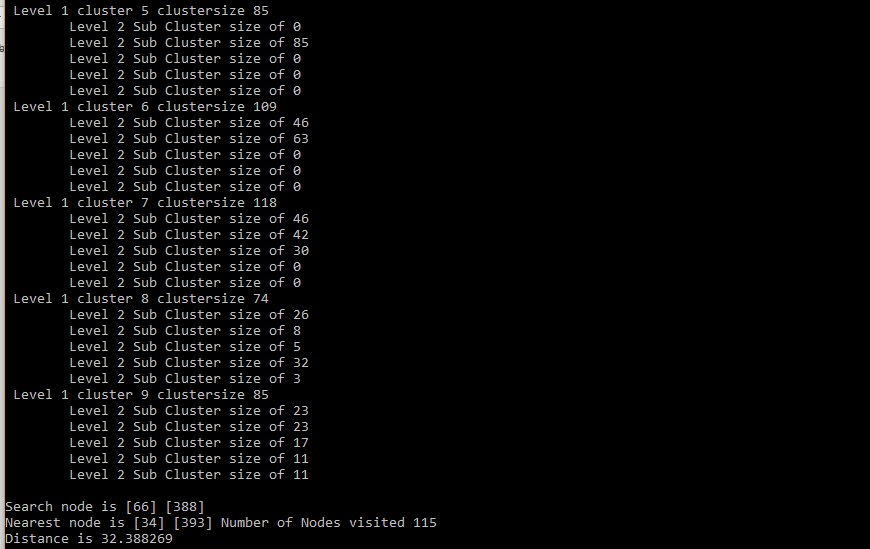
**How to execute**

"cs5331-HW4-(Arun kumar Jegarkal)-(Jaichandra Sesetty)-(Shivendra Satish Dubeer).exe"

Sample Execution

1000 Data Points with 2 Dimensions





10000 Data Points with 3 Dimensions

