

DATA MINING

Assignment 2

25-04-2014

Group Members:

Santhosh Alladi (2011A7PS086H)

Surya Nikhil R (2011A7PS201H)

Manideep Illendula (2011C6PS662H)

Observation:

Effect of number of clusters on Sum of Squared Errors (SSE):

As the number of cluster is changed, the SSE value is changed. As the number of clusters increases, the SSE value decreases.

Detailed description of the run information is given in the files attached.

Also we have done clustering using the two different formula for the calculation of intra cluster distances-Euclidean Distance and Manhattan Distance.

Effect of the change in the number of initial seeds:

First we set initial seed value to 10 and then we set the value to 5. We observed that the SSE value was comparatively less in the former case.