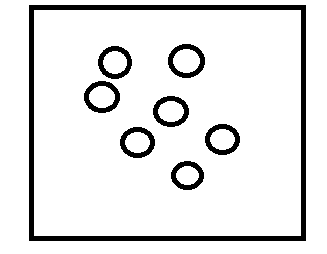
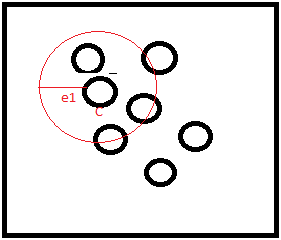
1. **[25] Prove that in DBSCAN, for a fixed MinPts (say M inP ts = 5) and two different neighborhood radius e1 < e2, a cluster C with respect to e1 must be a subset of a cluster C’ with respect to e2.**

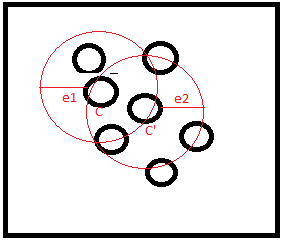
The definition of e-neighborhood can be defined with : Ne(C):{C’|d(C,C’)<=e), which describes objects that are within radius e from each other. Now if we take a look at some points:



We pick and arbitrary point and start building clusters such that, each cluster has at least 5 points, and each cluster has a distance of e. We can build our first cluster, C, with distance e1, and it will look something like this:



Now from here we can build another cluster, p, with distance e2, that satisfies our constraint d(C,C’)<=e, where the distance of C to C’ is less than e, in this case e1, and it will look something like this :



From here we can see that within a cluster C, if another cluster C’ is within e-neighborhood distance, the points of cluster C’ must be a subset and or contain points of cluster C.

**[25] Provide the pseudo code of the step in CLIQUE that finds dense cells in all sub-spaces**

(a) Partition each dimension into grids (subspaces)

(b) Let *k* = 1, repeat the following

(i). Travers all cells *c* in dimension *k*

1)if the number of points *p* in *c* is >= *j* , *j* being the density threshold, add cell *c* to a list of dense cells

(ii). Based on apriori, construct candidate sets from the (k+1)-dimensions

(c) Repeat until there are no more grids (subspaces) or candidate set is empty