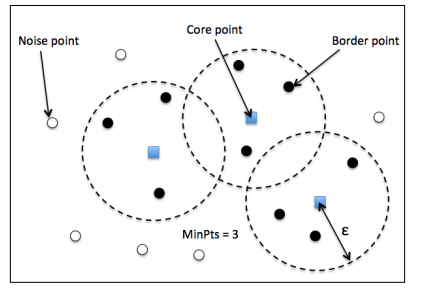
**Density-based Spatial Clustering of Applications with Noise**

The notion of density in DBSCAN is defined as the number of points within a specified radius 

In DBSCAN, special label is assigned to each sample (point) using the following criteria:

* A point is considered as **core point** if at least a specified number (MinPts) of neighboring points fall within the specified radius 
* A **border point** is a point that has fewer neighbors than MinPts within ,but lies within the  radius of a core point.
* All other points that are neither core nor border points are considered as **noise points**.

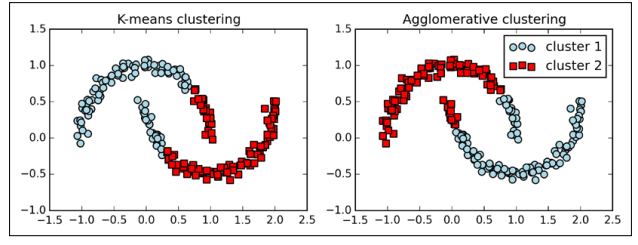


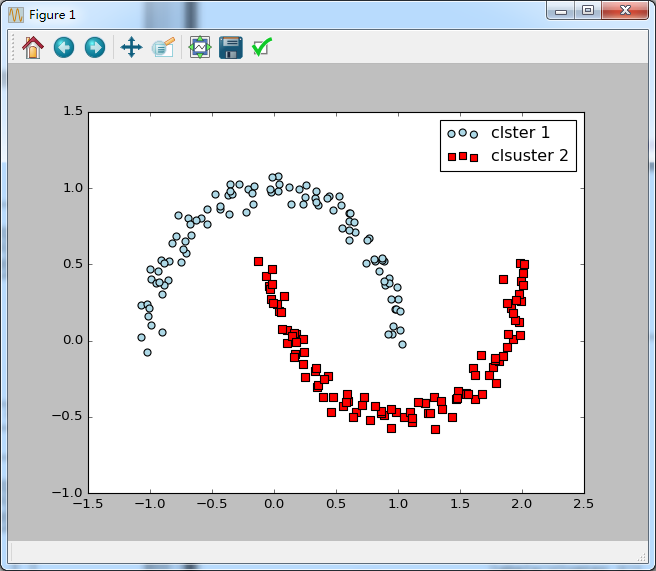
Two simples steps:

1. Form a separate cluster for each core point or a connected group of core points (core points are connected if they are no farther away than )
2. Assign each border point to the cluster of its corresponding core point.

Advantages:

1. It does not assume that the clusters have a spherical shape as in k-means.
2. It doesn’t necessarily assign each point to a cluster but is capable of removing noise points





Disadvantages:

1. With an increasing number of features in our dataset—given a fxed size training set—the negative effect of the curse of dimensionality increases.
2. In addition, we have two hyperparameters in DBSCAN (MinPts and ε ) that need to be optimized to yield good clustering results. Finding a good combination of MinPts and ε can be problematic if the density differences in the dataset are relatively large.