

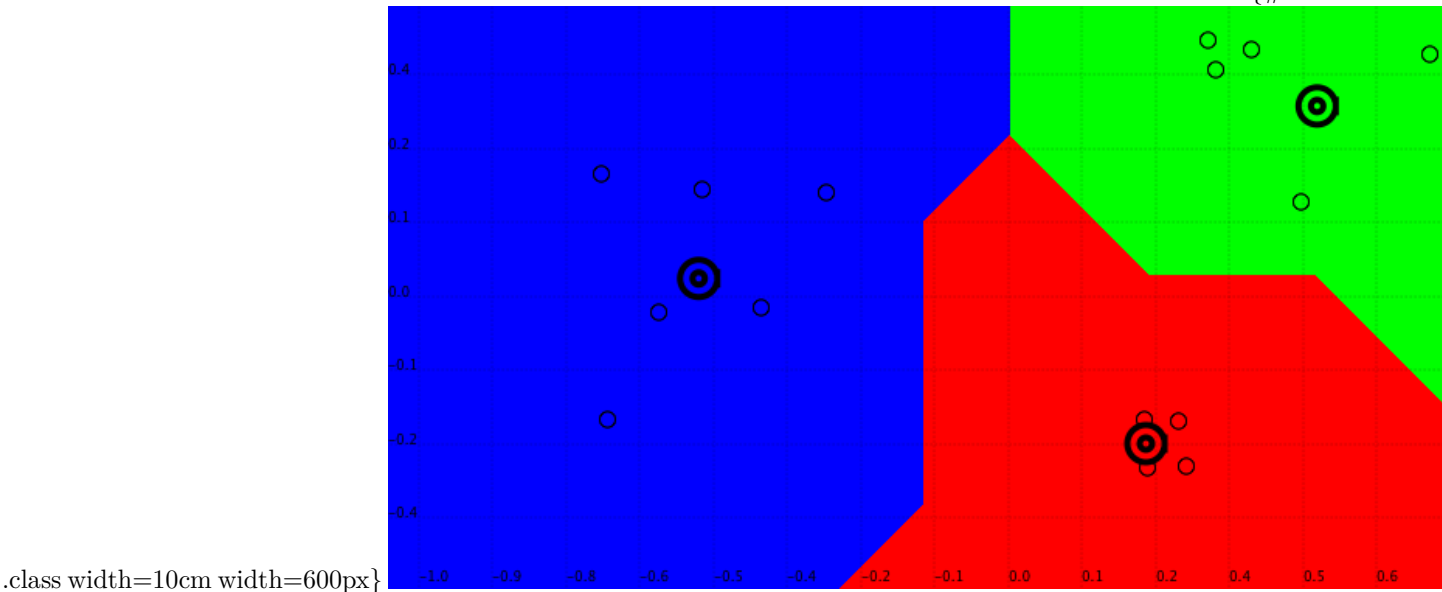
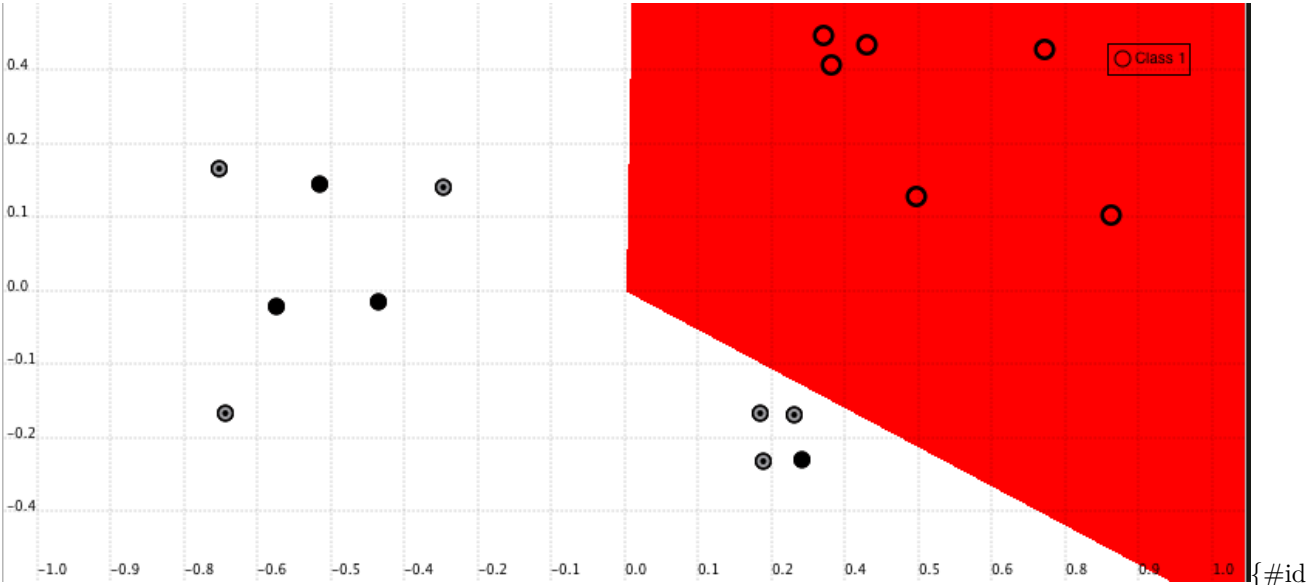
Data Mining Homework 7

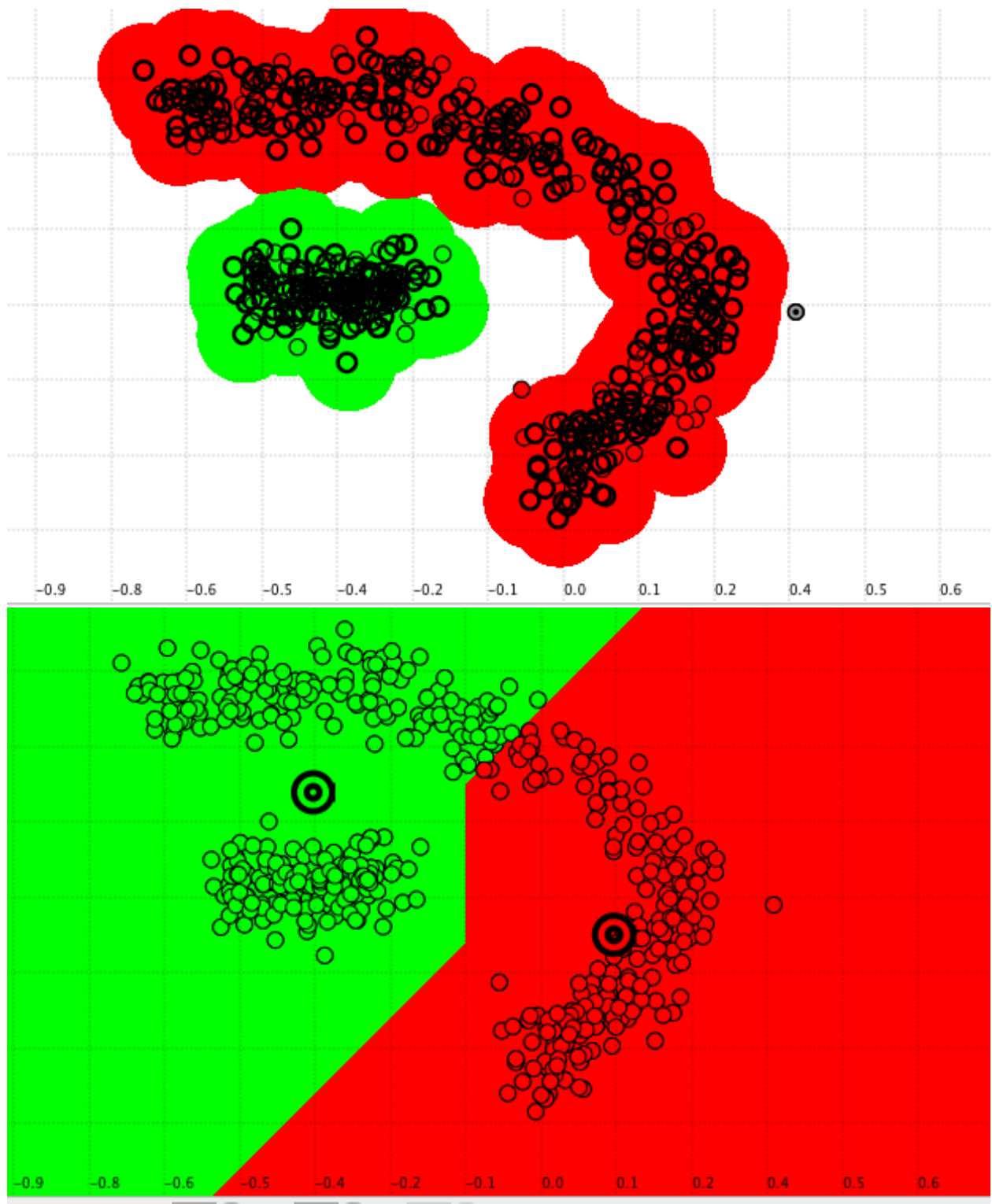
Madis Nõmme

April 3rd, 2014

1 DBSCAN vs K-means

1.1 K-means





1.1.1 Pros

1. Easy to understand and implement
2. Fast

1.1.2 Cons

1. Can't find arbitrary shaped clusters
2. One has to know the number of clusters
3. Choosing initial cluster centers affects results

1.1.3 DBSCAN

1.1.4 Pros

1. No need to know cluster centers

1.1.5 Cons

1. Can't find clusters with highly varying density
2. Requires