

# Study and visualiser of different clustering methods for density-based clustering

Martin Loginov, Hans Mäesalu, Sven  
Aller



# Main goals

- Study different density-based clustering methods
  - DBSCAN
  - DCBOR
  - SNN
- Implement a clustering visualisation application



# DBSCAN

Density-based spatial clustering of applications with noise

- Based on density based notion of clusters
- Can discover clusters of different shapes
- Has notion of noise
- Do not need to know the number of clusters beforehand

# DCBOR

## Density Clustering Based on Outlier Removal

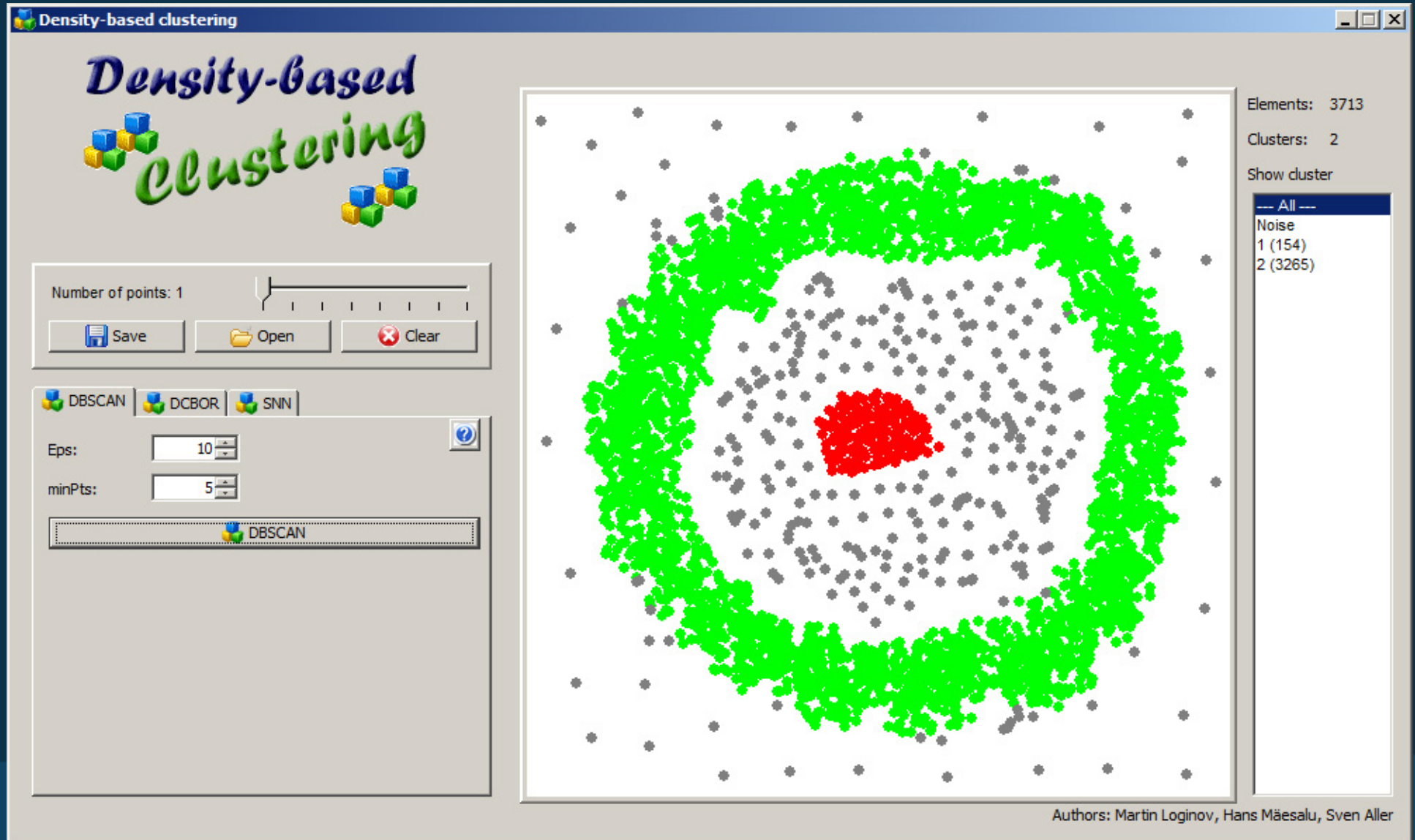
- Based on the notion of local density
  - needs to find k-nearest neighbors of each datapoint
  - we use a KD-tree
- Two phases
  - Outlier removal
    - Only 1 input parameter!
  - Clustering
    - cross between DBSCAN and single-link

# SNN

## Shared nearest neighbor

- Similarity definition based on the number of shared nearest neighbors
  - If point  $p_1$  is close to point  $p_2$  and they are both close to a set of points,  $S$ , then their similarity is equal to the number of points in set  $S$
- Density is defined as the number of points that are similar to a point

# Implementation





Thank you!