

Recap 5

Dimensionality Reduction & Clustering

ISLR 12, ESL 14, tSNE



Jilles Vreeken
Krikamol Muandet



UNIVERSITÄT
DES
SAARLANDES



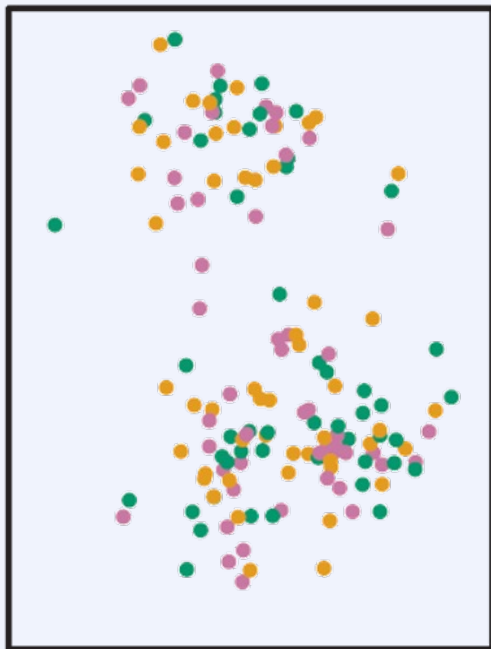
CISPA
HELMHOLTZ CENTER FOR
INFORMATION SECURITY

Lecture Recap 2

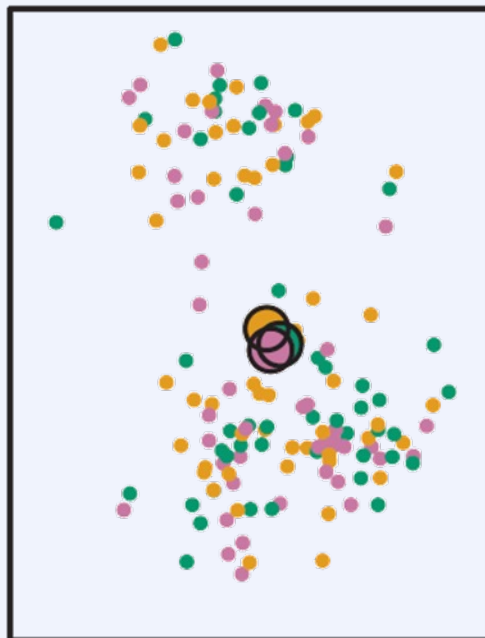
- K-Means
 - Clustering: group similar points into *clusters*
 - K-means: find clusters defined through a centroid and assign closest points to it
 - Efficient algorithm with convergence to local optimum

K-Means

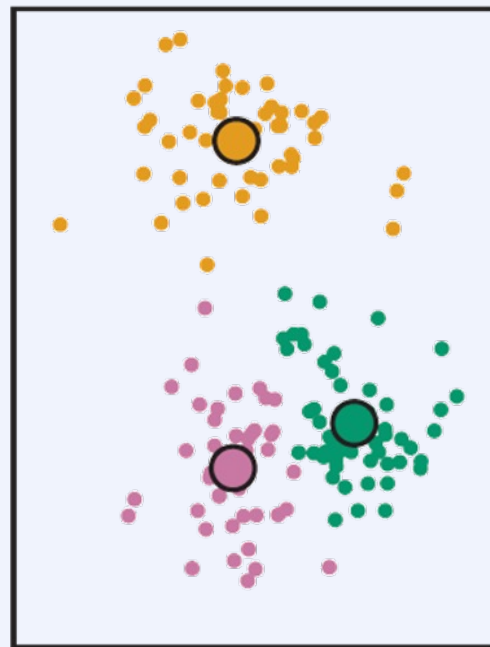
Step 1



Iteration 1, Step 2a



Final Results

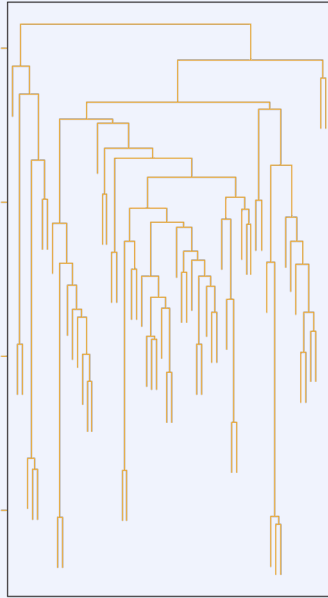


Lecture Recap 2

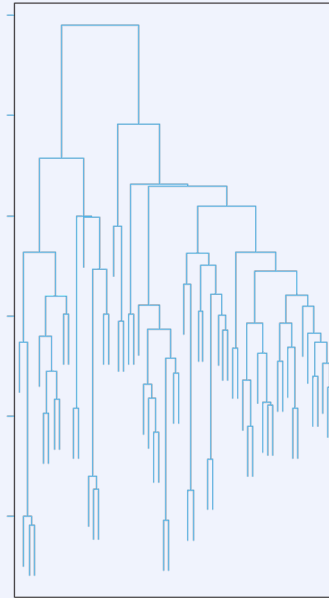
- K-Means
 - Clustering: group similar points into *clusters*
 - K-means: find clusters defined through a centroid and assign closest points to it
 - Efficient algorithm with convergence to local optimum
- Hierarchical Clustering
 - Iteratively merge closest clusters bottom-up/top-down
 - Dendrogram shows hierarchy of clusters
 - Different linkage (single, complete) results in vastly different results

Hierarchical Clustering

Average Linkage



Complete Linkage



Single Linkage

