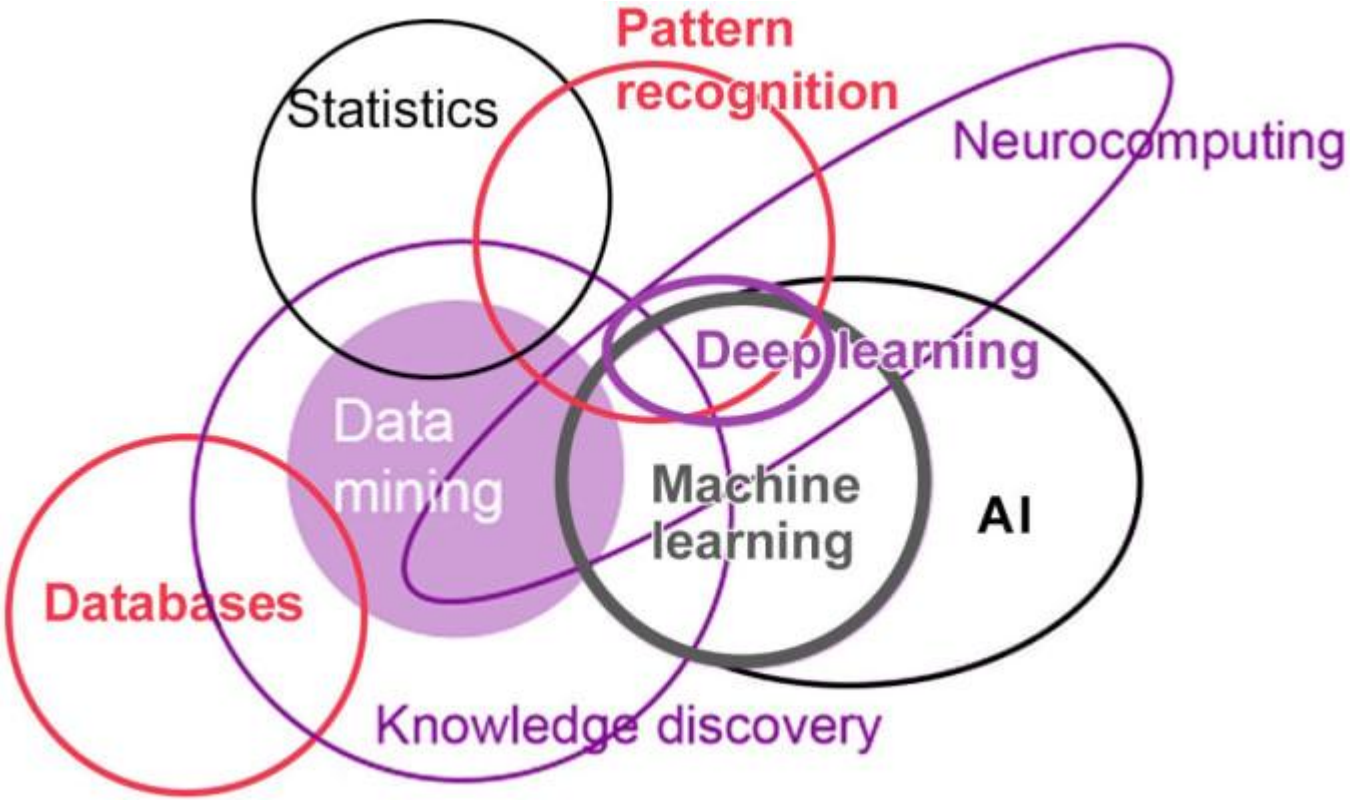
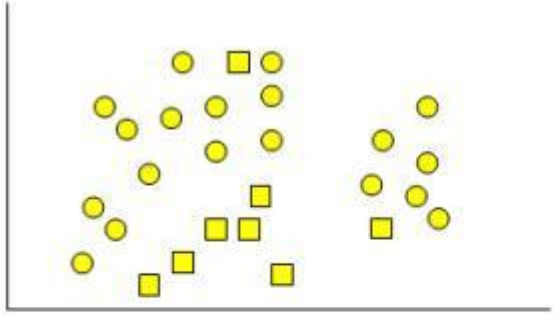
MIDS W207
Applied Machine Learning

Week 9
Live Session Slides

Data Mining and Machine Learning

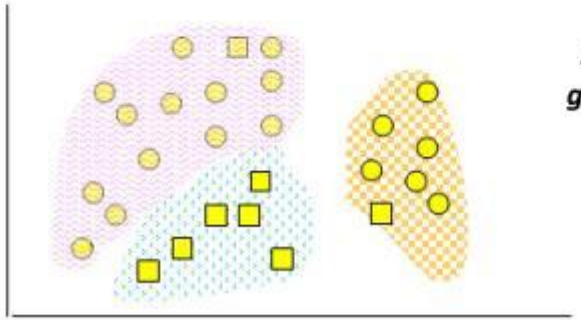


Clustering vs Classification



Classification

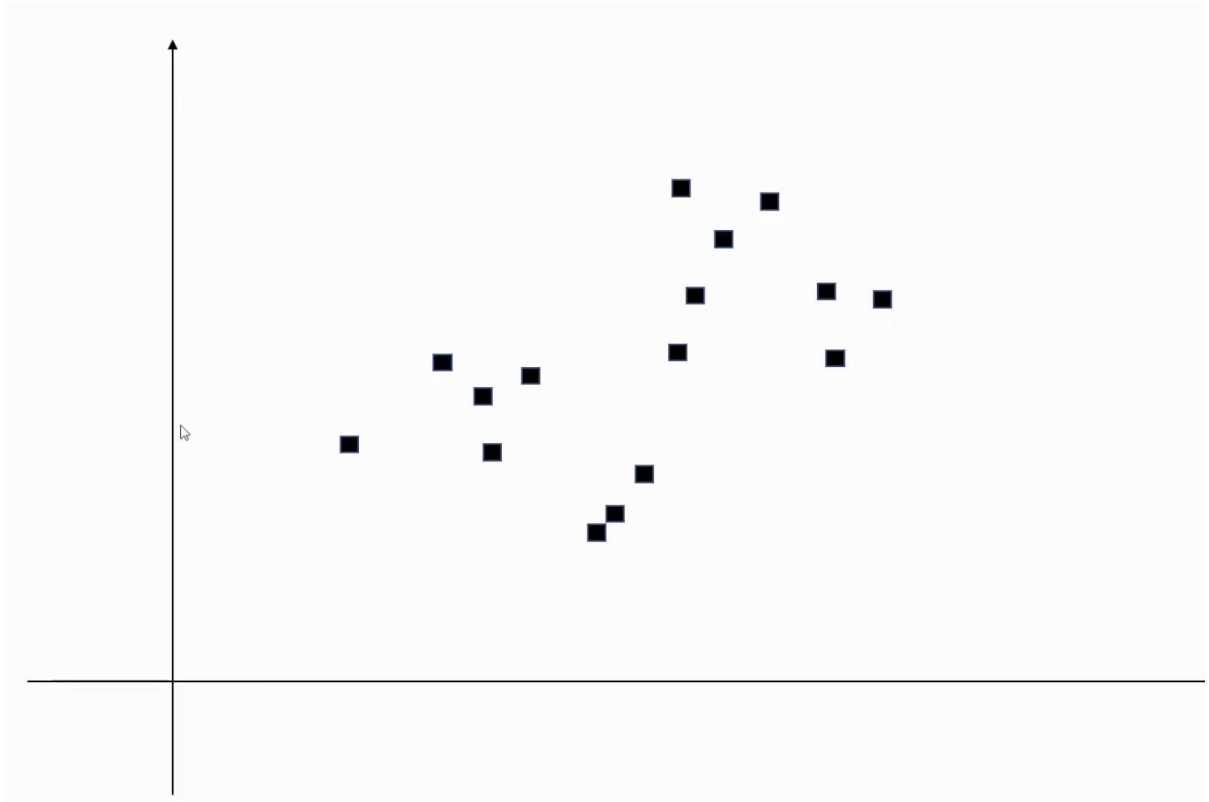
Classes are defined before data processing.



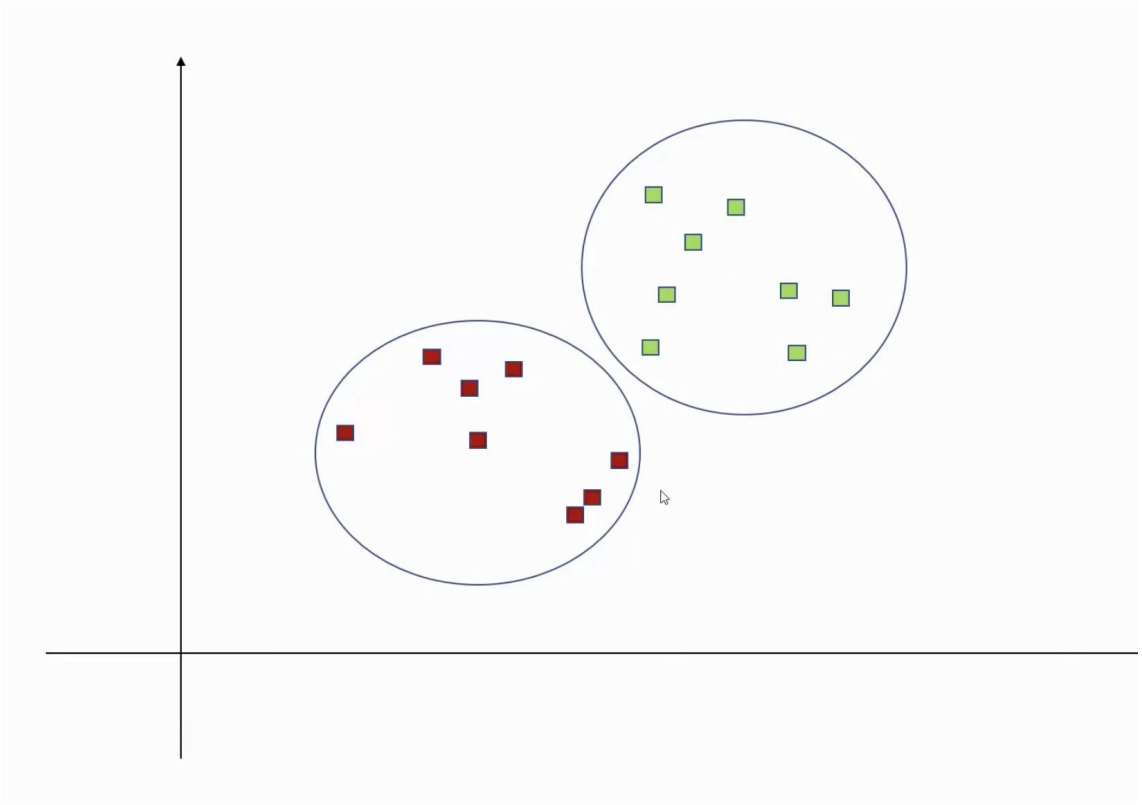
Clustering

Classes are not defined beforehand. Data mining searches for homogeneous groups, groups of objects that have common properties.

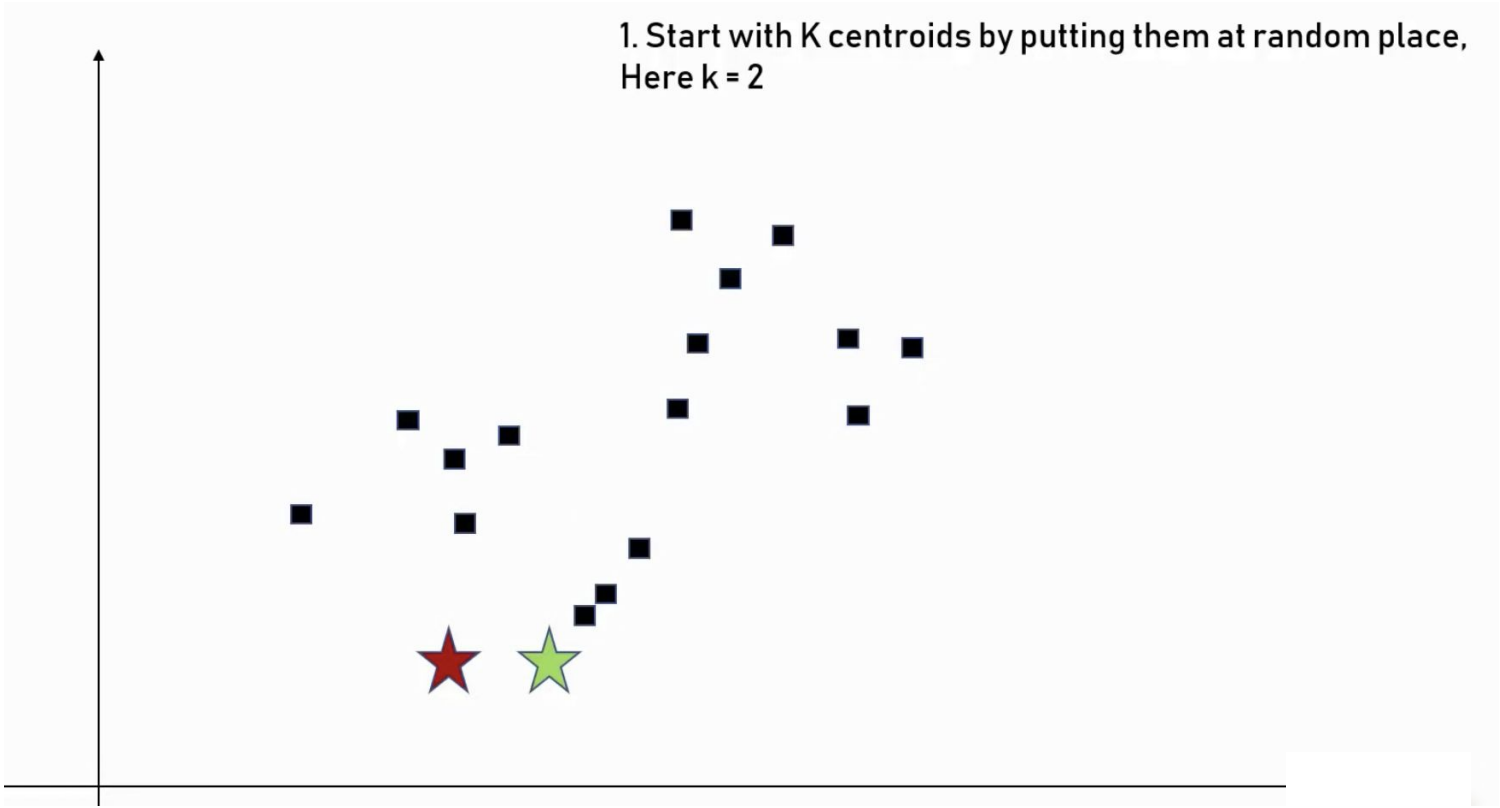
K-Means Clustering



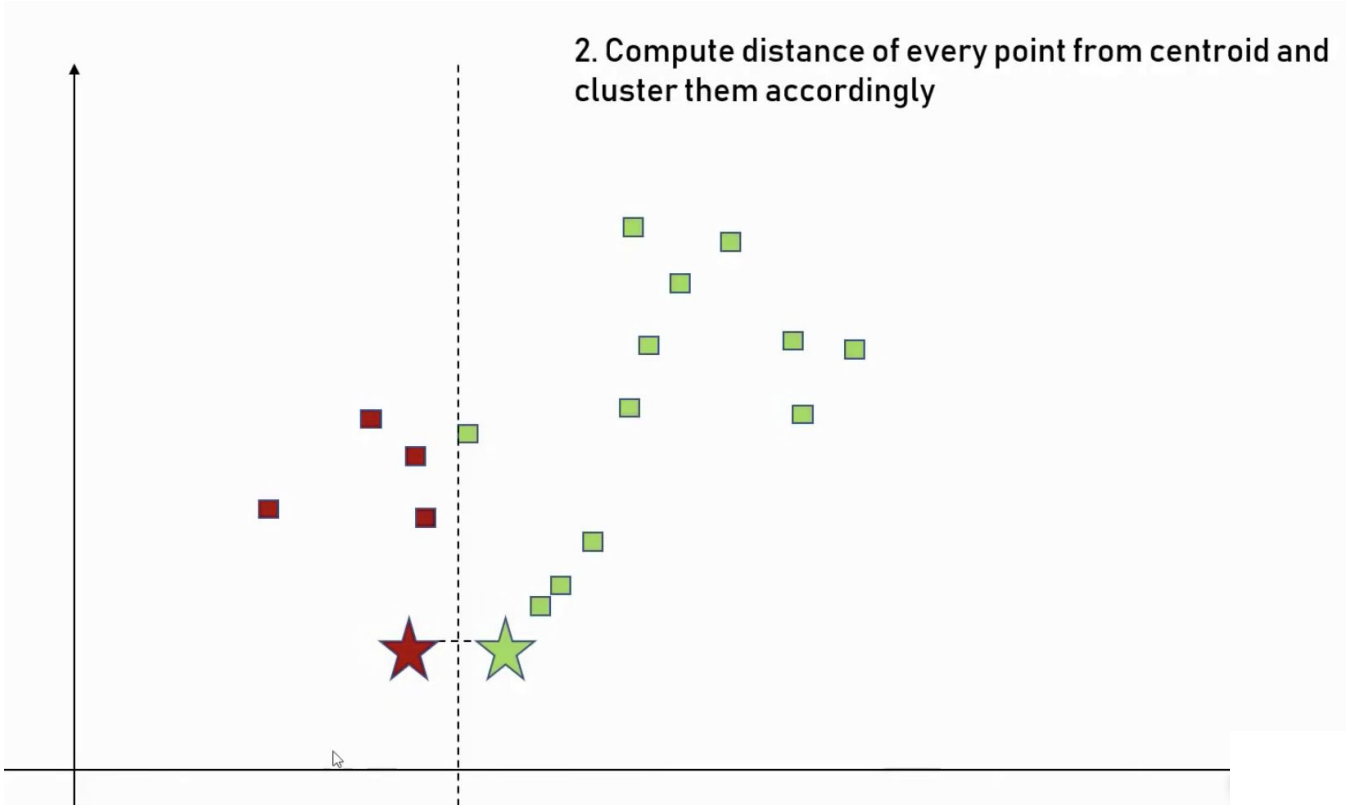
K-Means Clustering



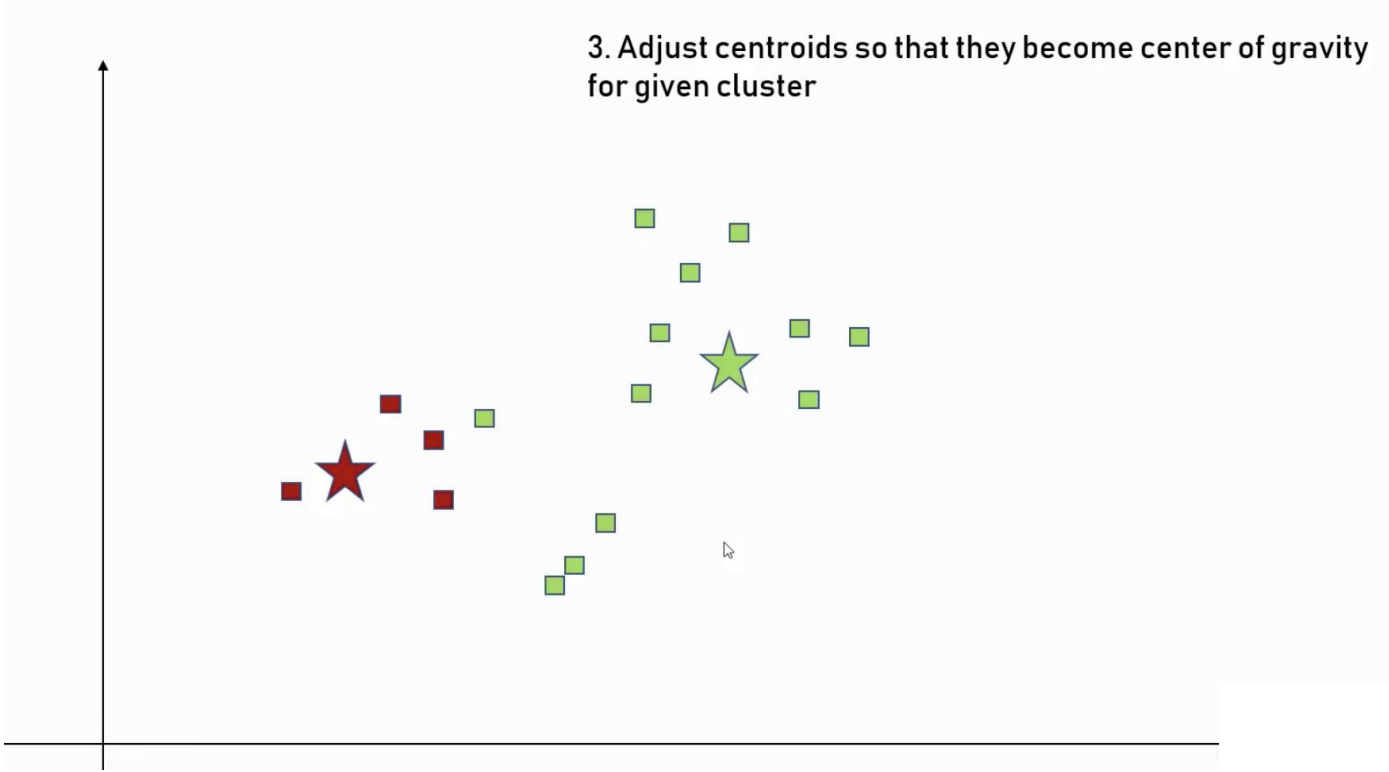
K-Means Clustering



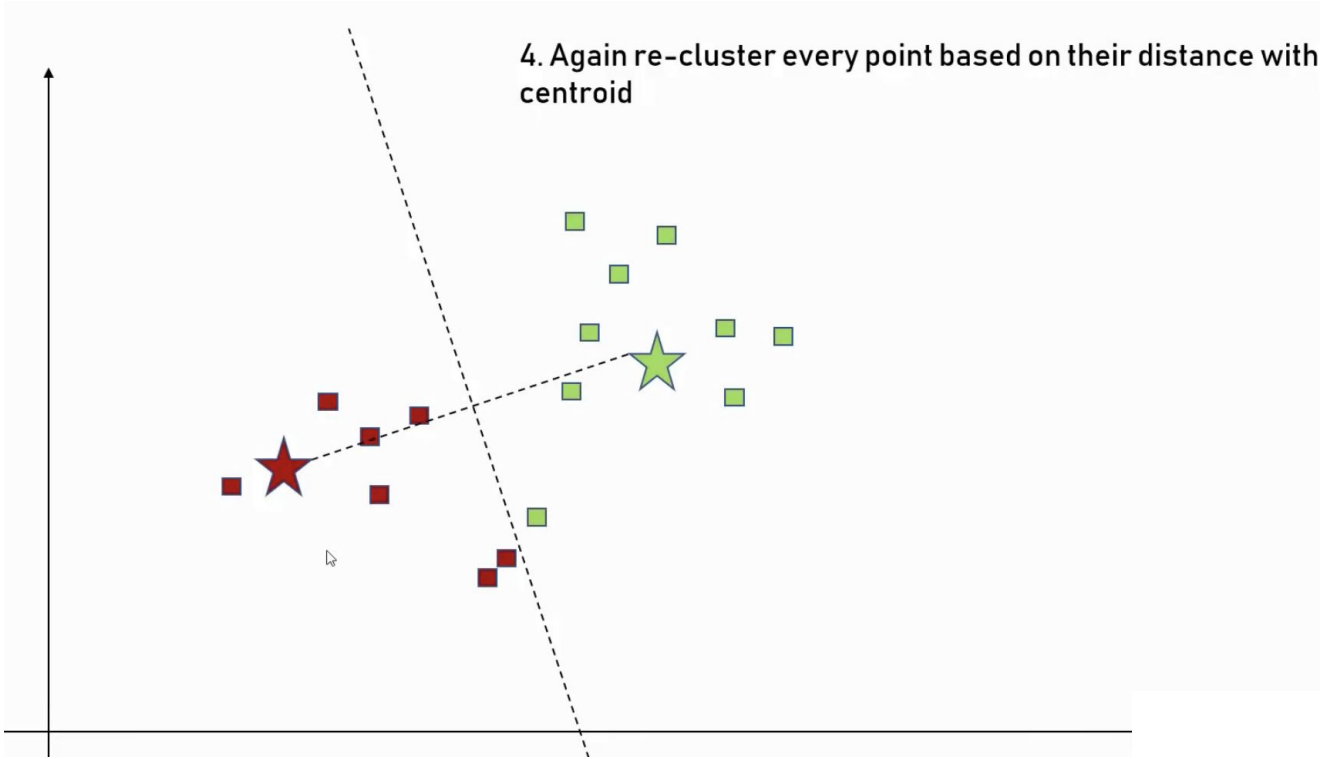
K-Means Clustering



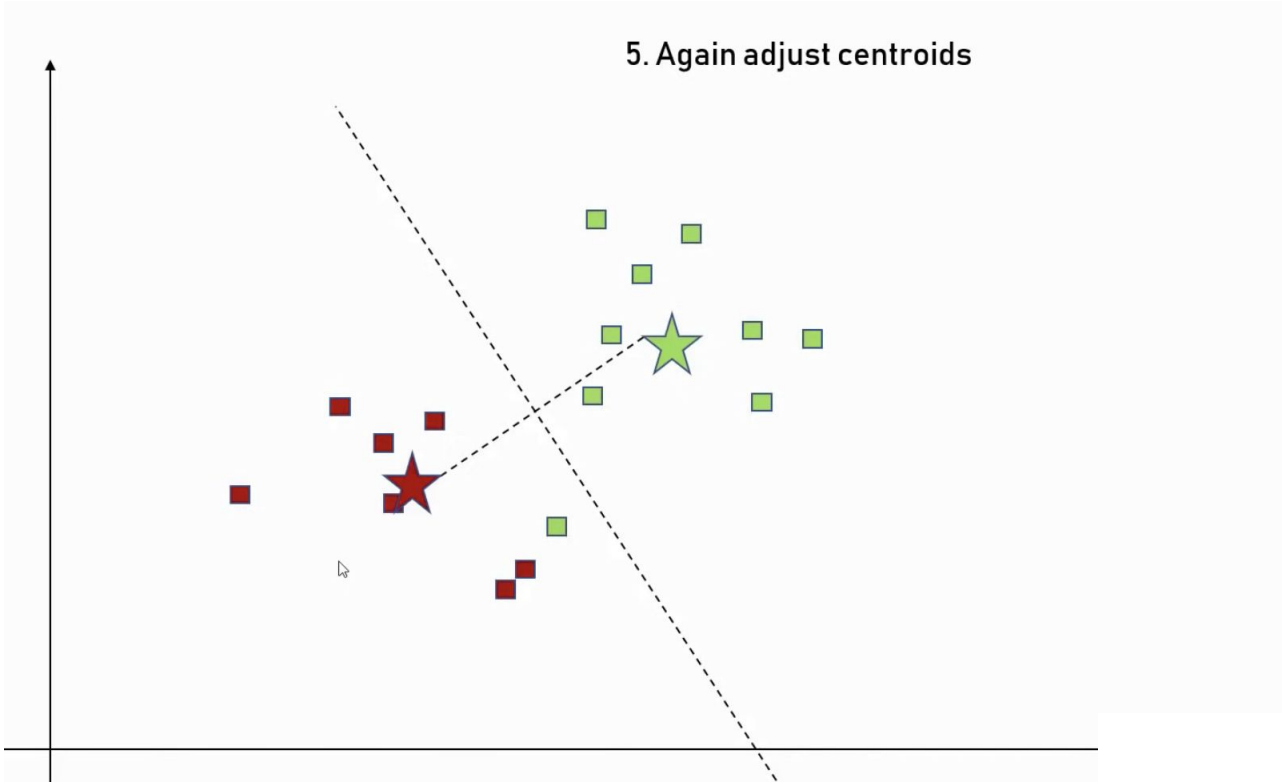
K-Means Clustering



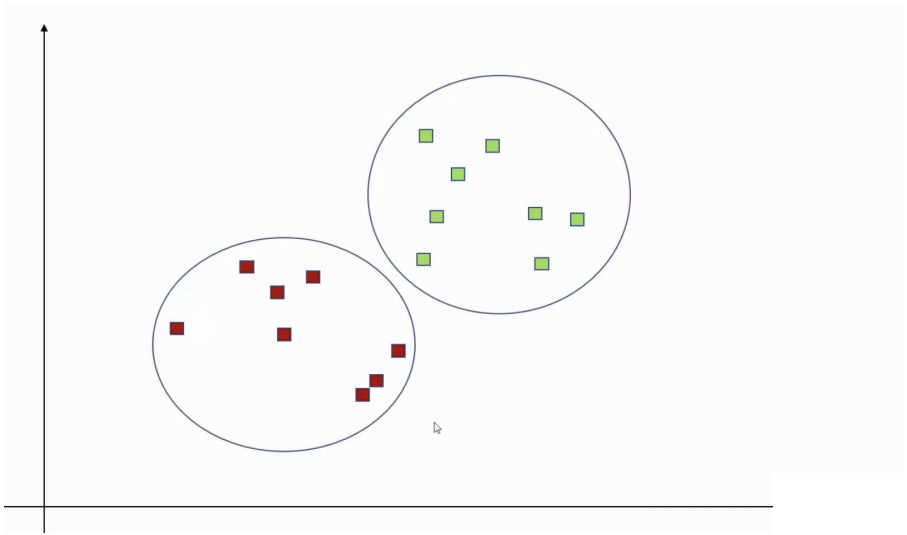
K-Means Clustering

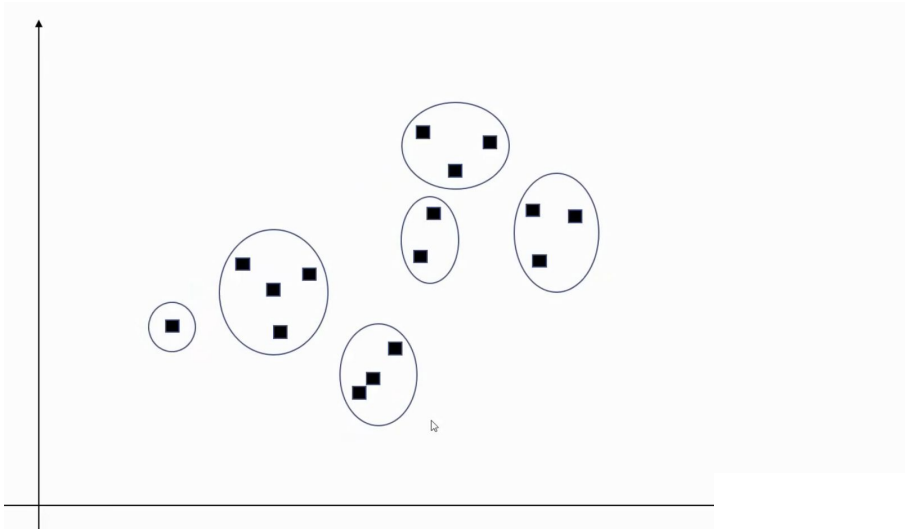


K-Means Clustering

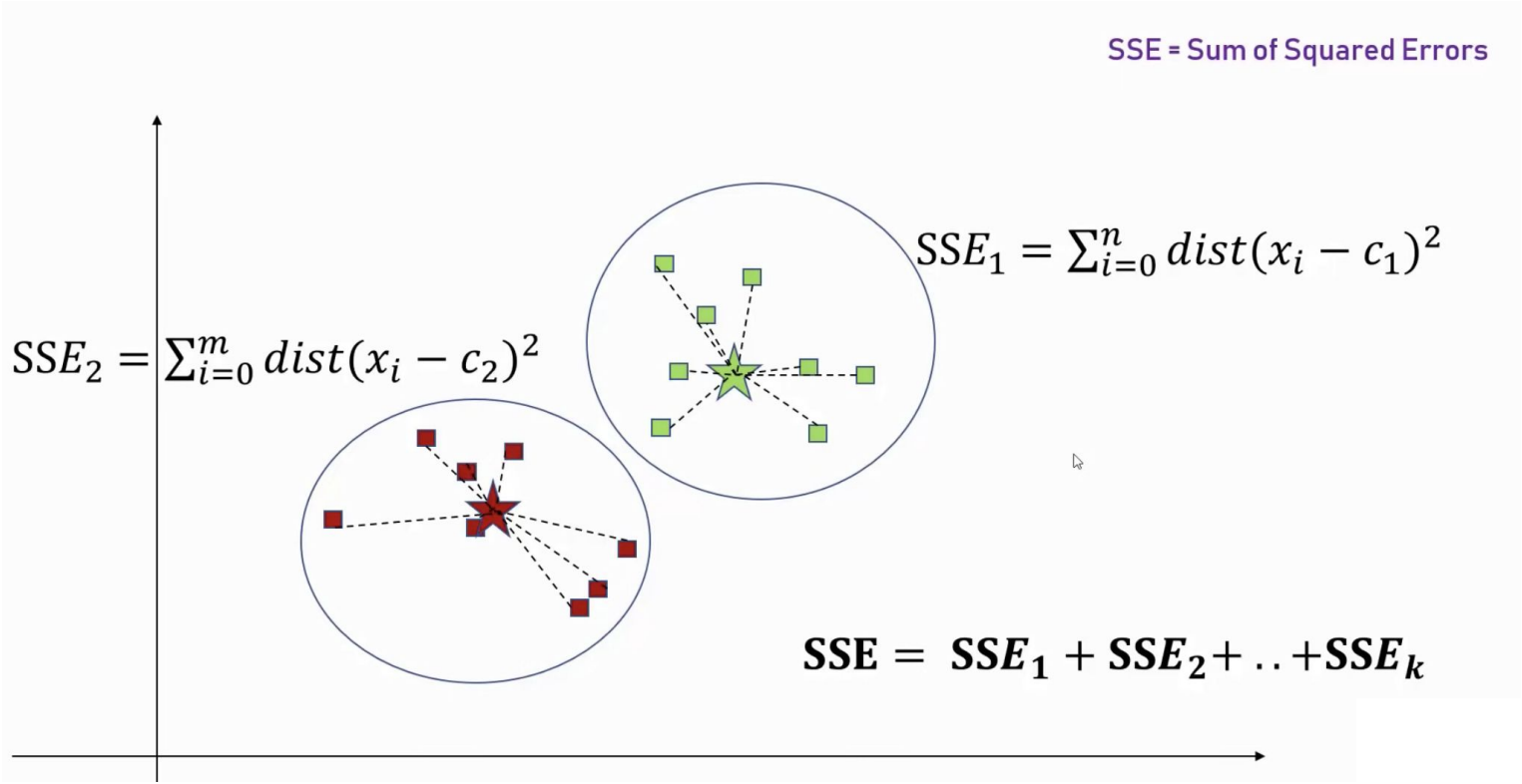


K-Means Clustering: Finding k

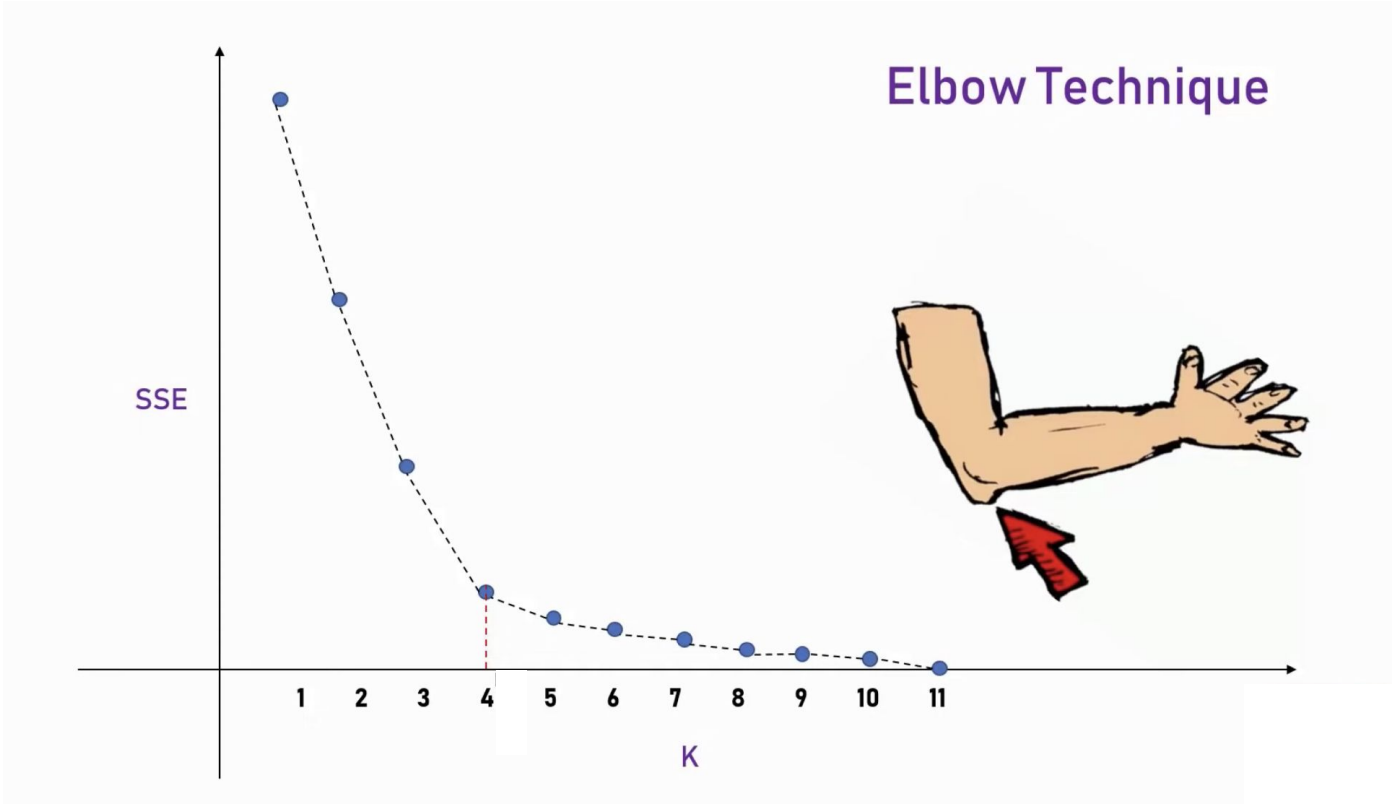




K-Means Clustering: Finding k



K-Means Clustering: Finding k

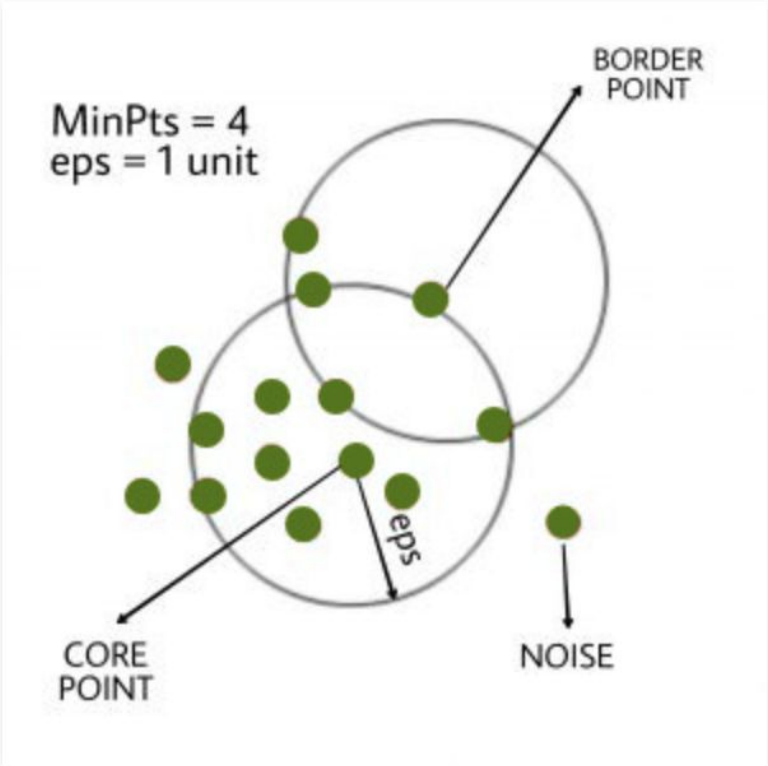


K-Means Algorithm

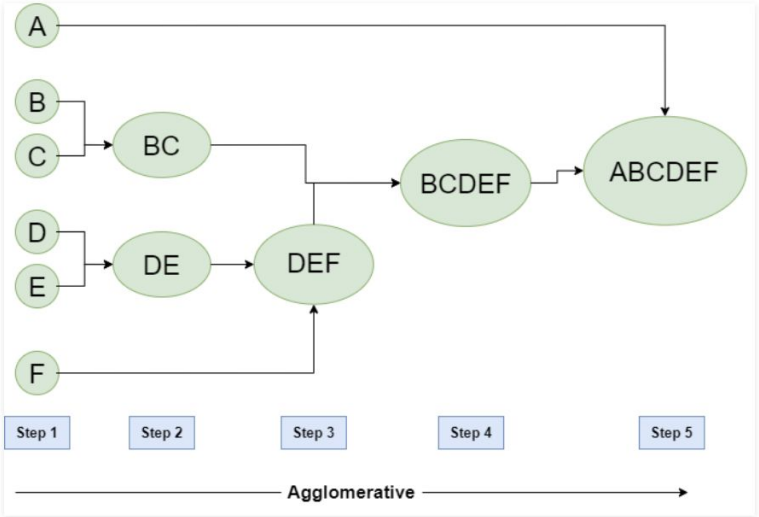
Algorithm 1 k-means algorithm

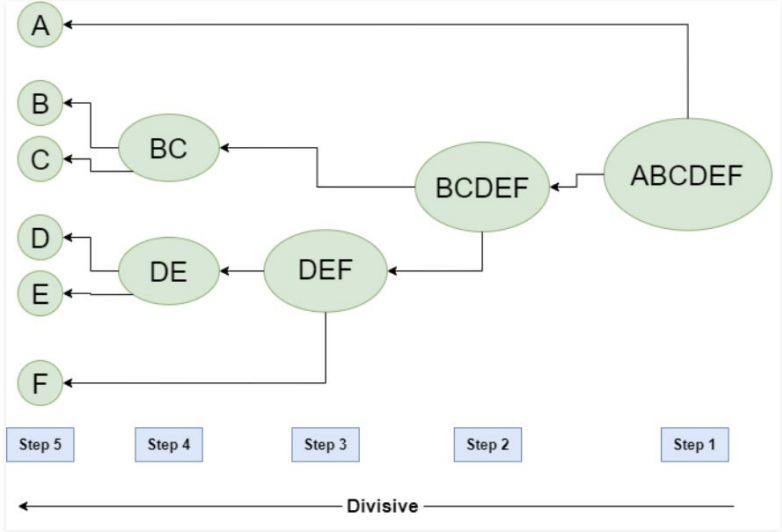
1: Specify the number k of clusters to assign.
2: Randomly initialize k centroids.
3: **repeat**
4: **expectation:** Assign each point to its closest centroid.
5: **maximization:** Compute the new centroid (mean) of each cluster.
6: **until** The centroid positions do not change.

K-Means Algorithm: Types: DBSCAN



K-Means Algorithm: Types: Hierarchical





Code Review

Breakout Session Exercise