

K-means Method for Clustering

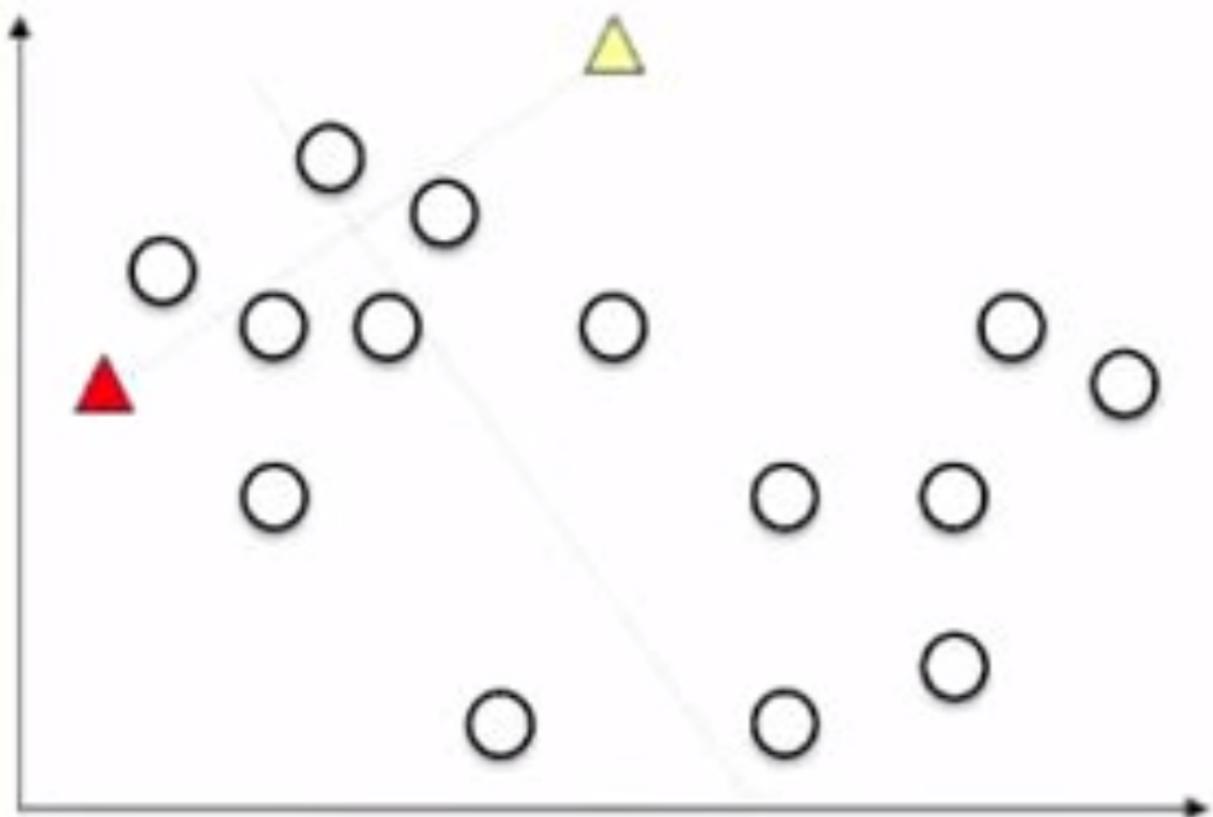
(https://youtu.be/_aWzGGNrcic)

K-means clustering algorithm

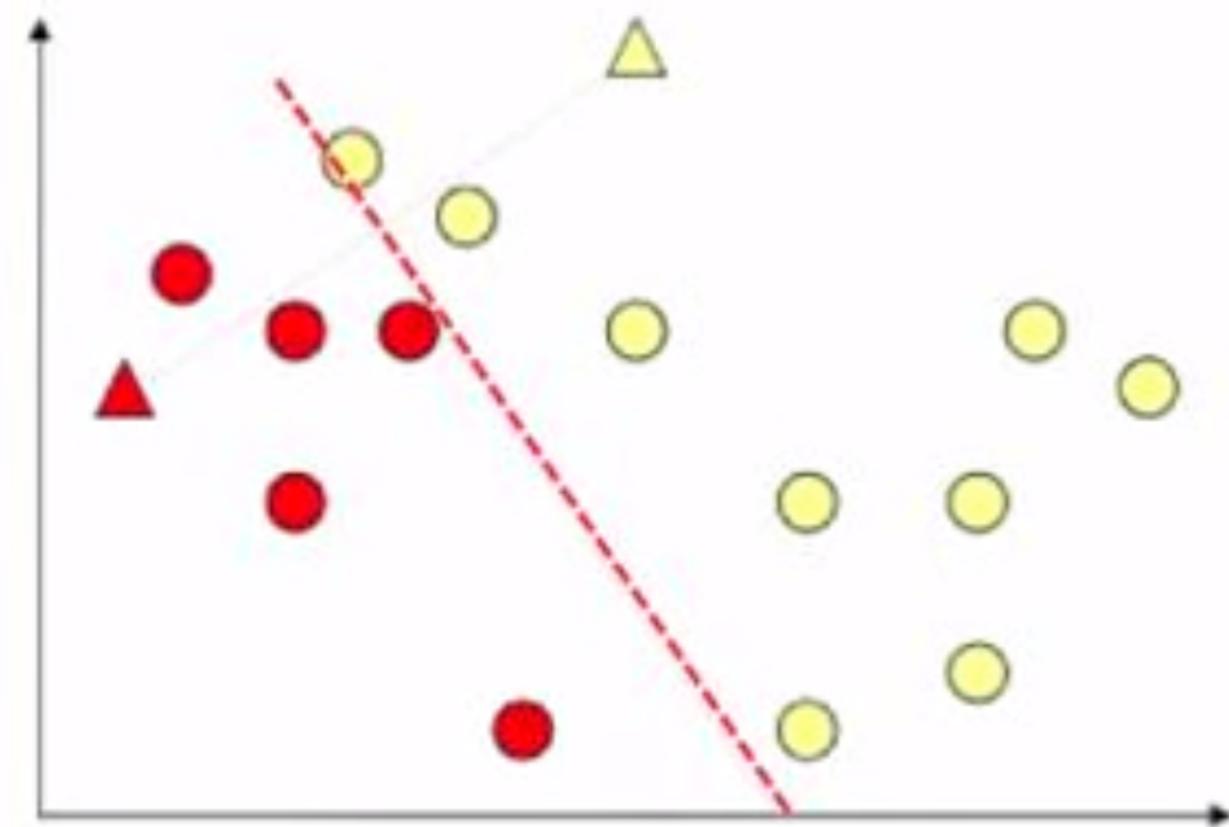
- Input: K, set of points $x_1 \dots x_n$
 - Place centroids $c_1 \dots c_K$ at random locations
 - Repeat until convergence:
 - for each point x_i :
 - find nearest centroid c_j $\arg \min_j D(x_i, c_j)$
 - assign the point x_i to cluster j
 - for each cluster $j = 1 \dots K$:
$$c_j(a) = \frac{1}{n_{j, x_i \rightarrow c_j}} \sum x_i(a) \quad \text{for } a = 1 \dots d$$
 - new centroid c_j = mean of all points x_i assigned to cluster j in previous step
 - Stop when none of the cluster assignments change
- $O(\#iterations * \#clusters * \#instances * \#dimensions)$

(<https://youtu.be/aWzGGNrcic>)

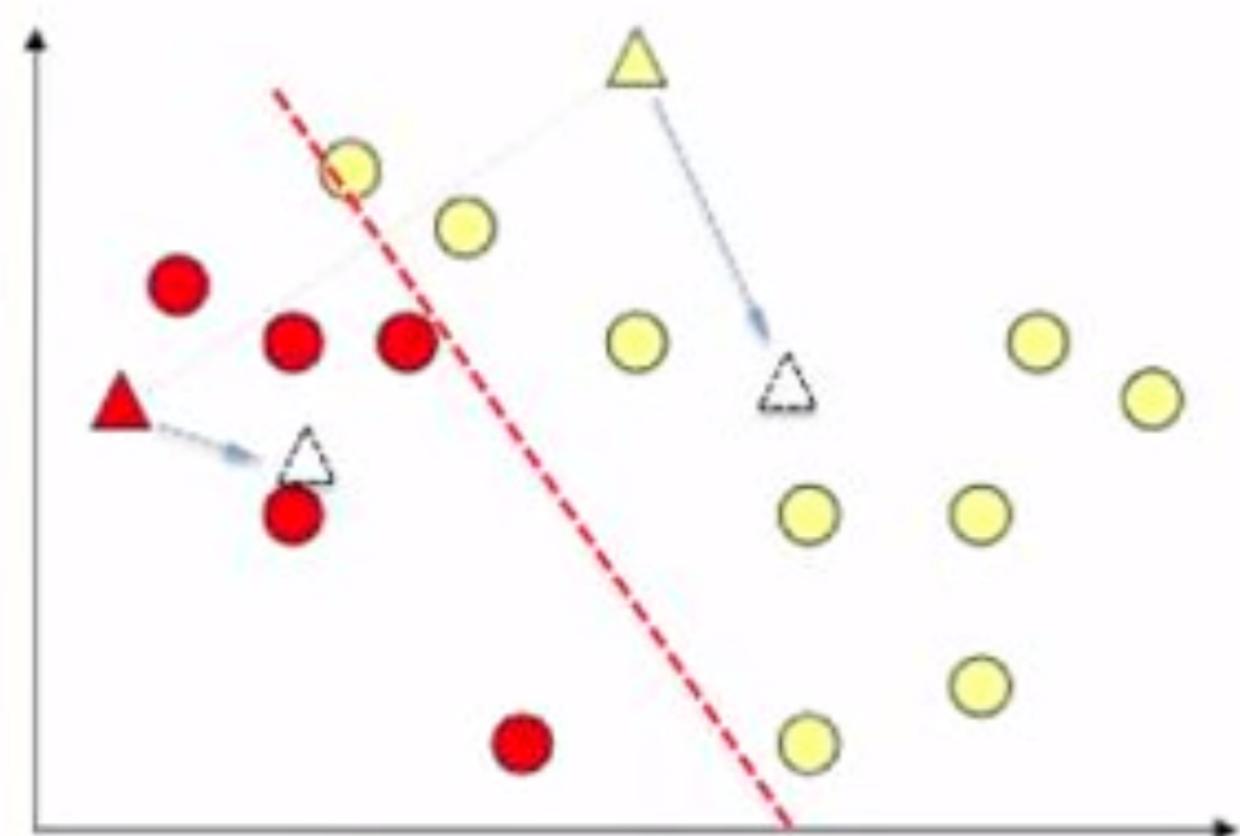
K-means clustering example



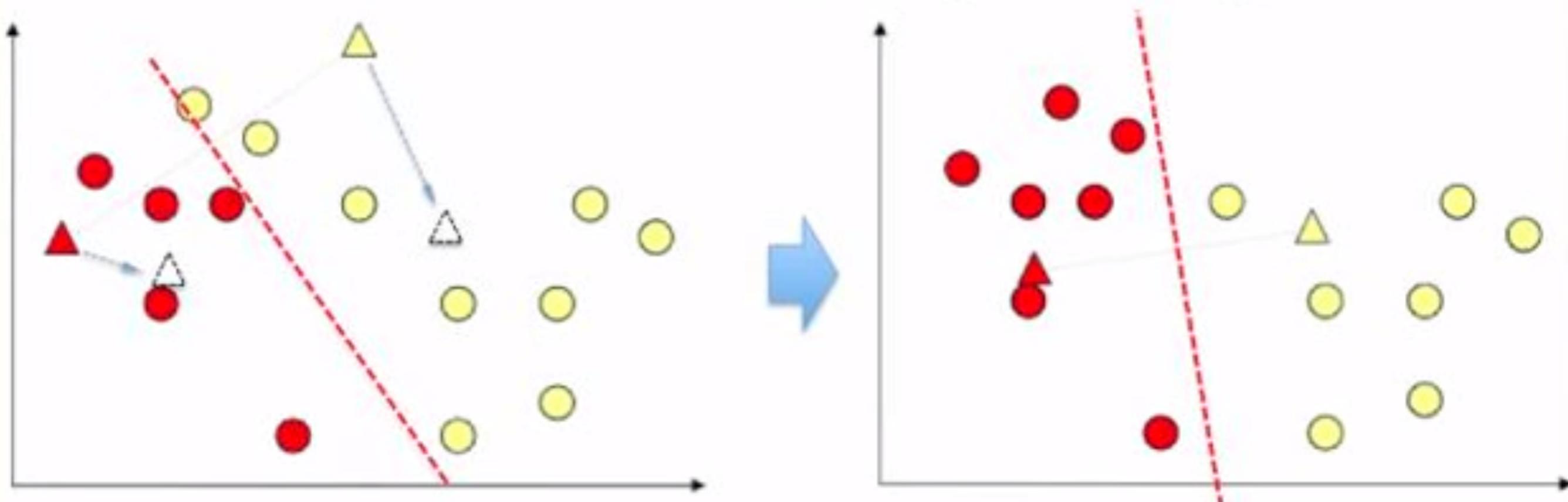
K-means clustering example



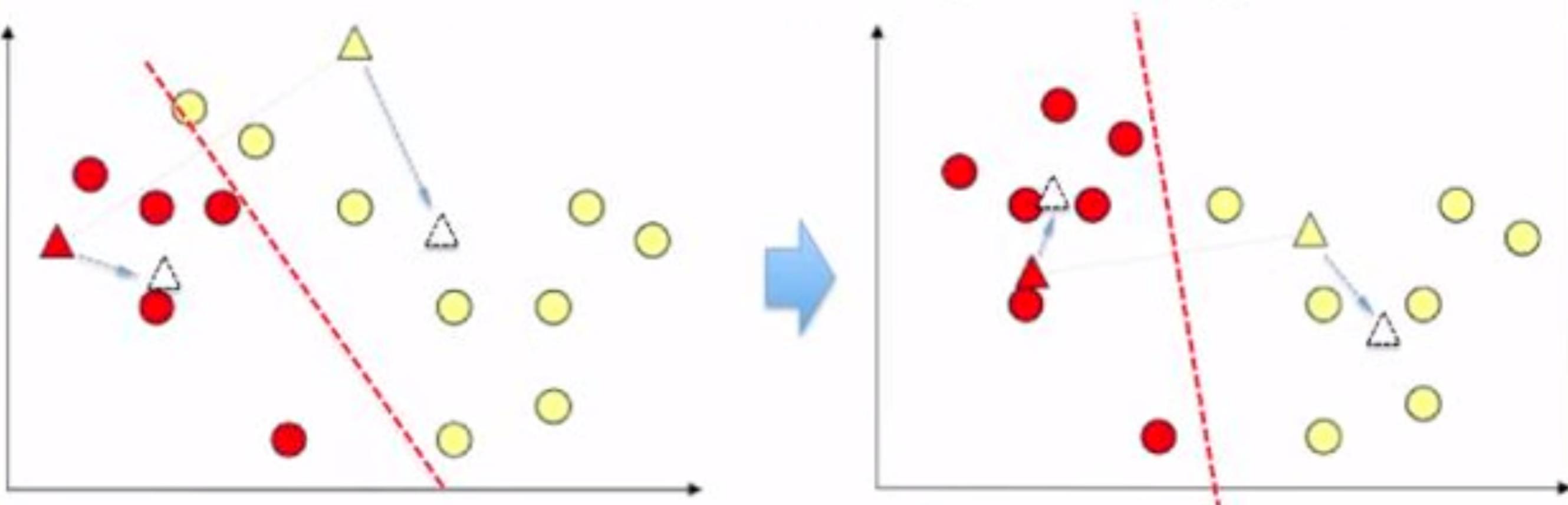
K-means clustering example



K-means clustering example



K-means clustering example



K-means clustering example

