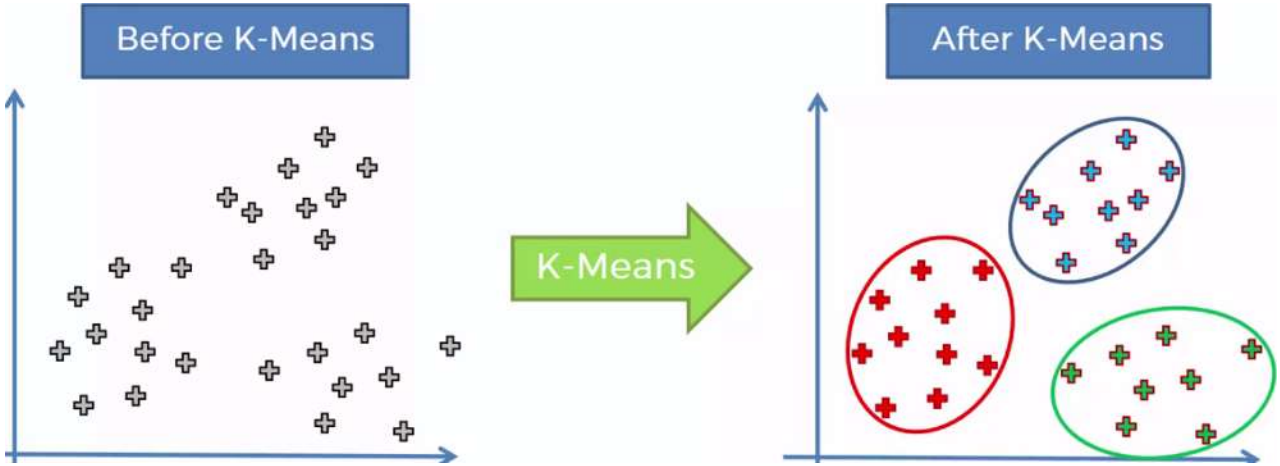
Unsupervised Learning

Clustering technique: k-means

Clustering technique: k-means

- k-means is one of the simplest unsupervised learning algorithms
- Solves the well known clustering problem
- Used when you have unlabeled data
- The goal is to find groups in the data, with the number of groups represented by the variable K.

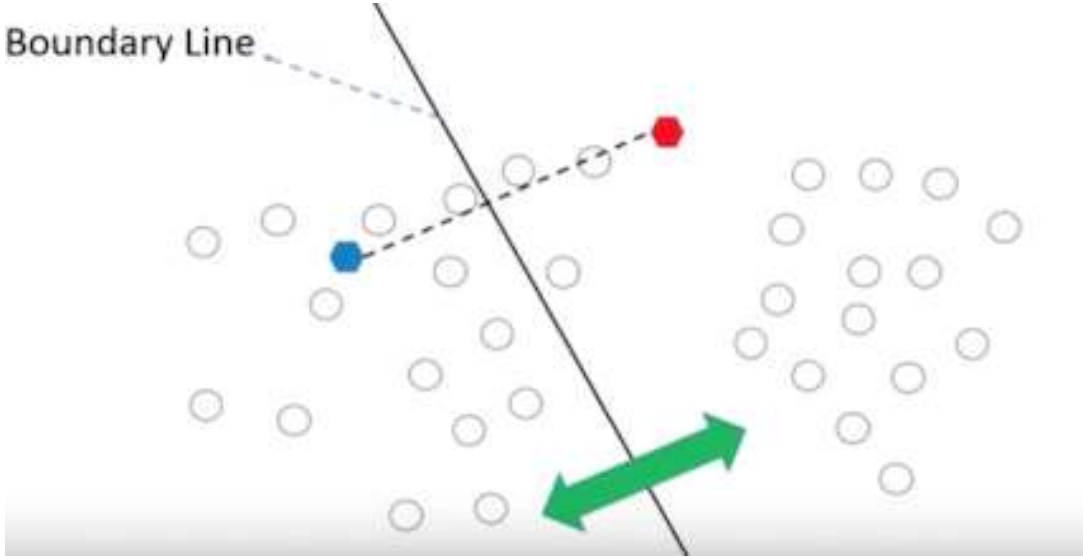


How it works – Step 1: Initialization

- Specify the desired number of clusters K and randomly choose K data points from the dataset as initial centroids.

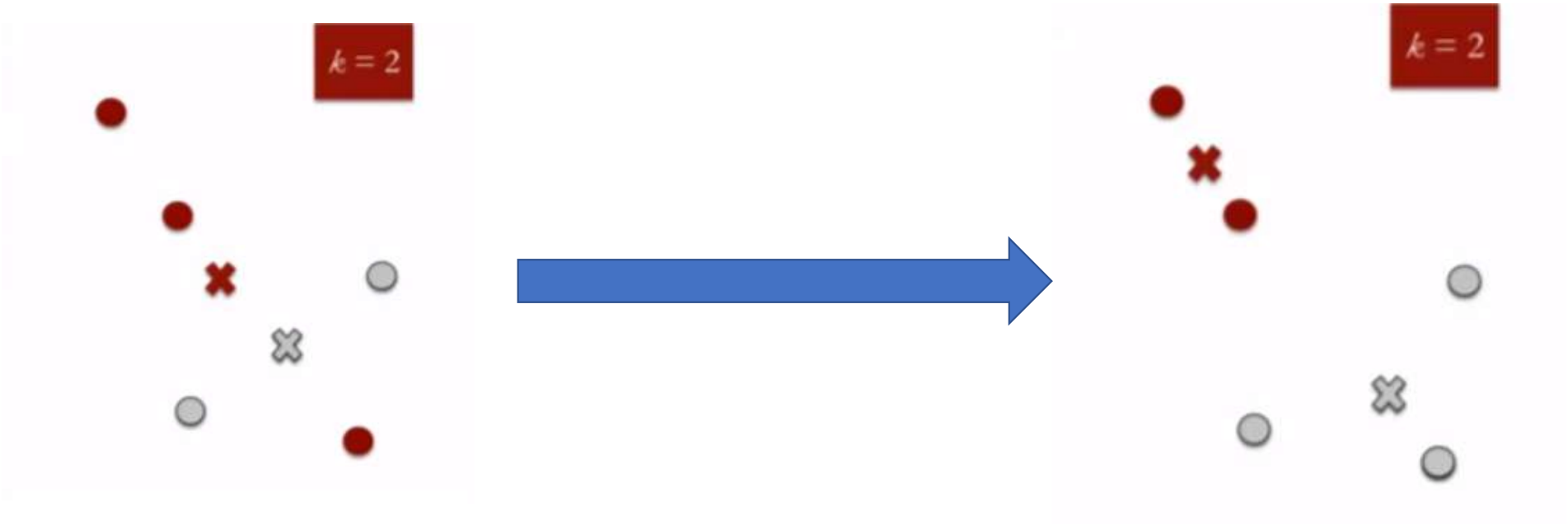
How it works – Step 2: Cluster Assignment

- Assign each data point to a cluster.
- All the data points that are the closest to a centroid will create a cluster.



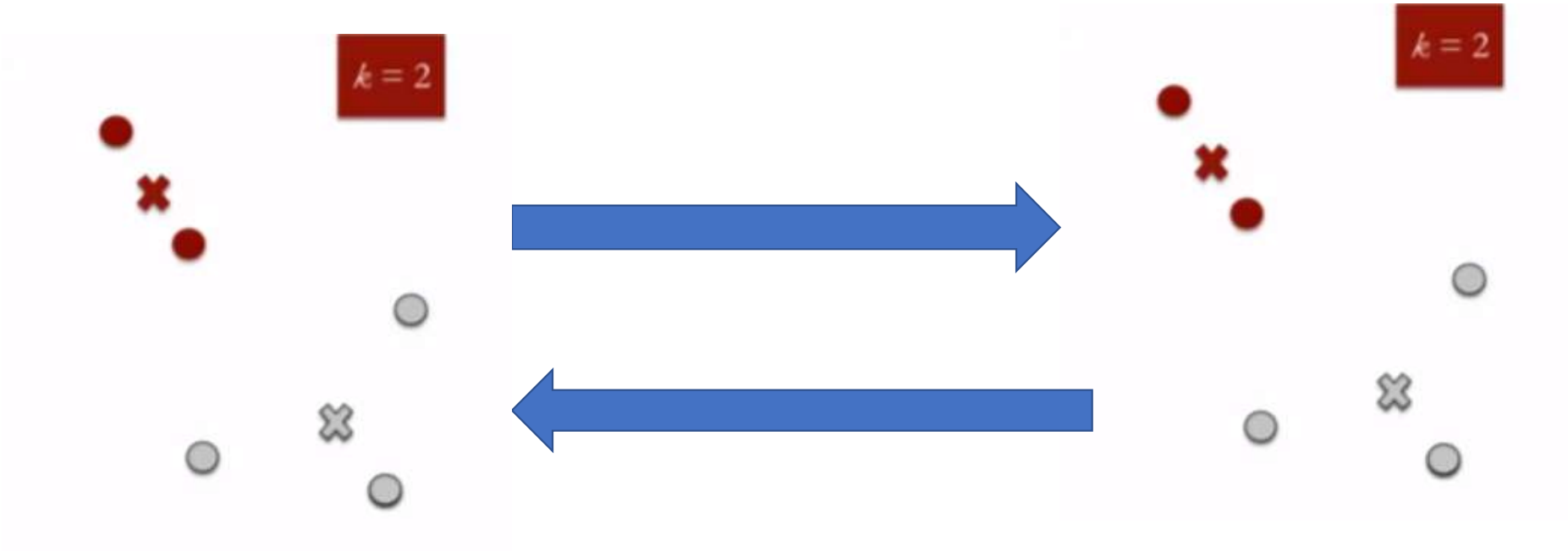
How it works – Step 3: Move the centroid

- Now, we have new clusters, that need centers. A centroid's new value is going to be the mean of all the examples in a cluster.

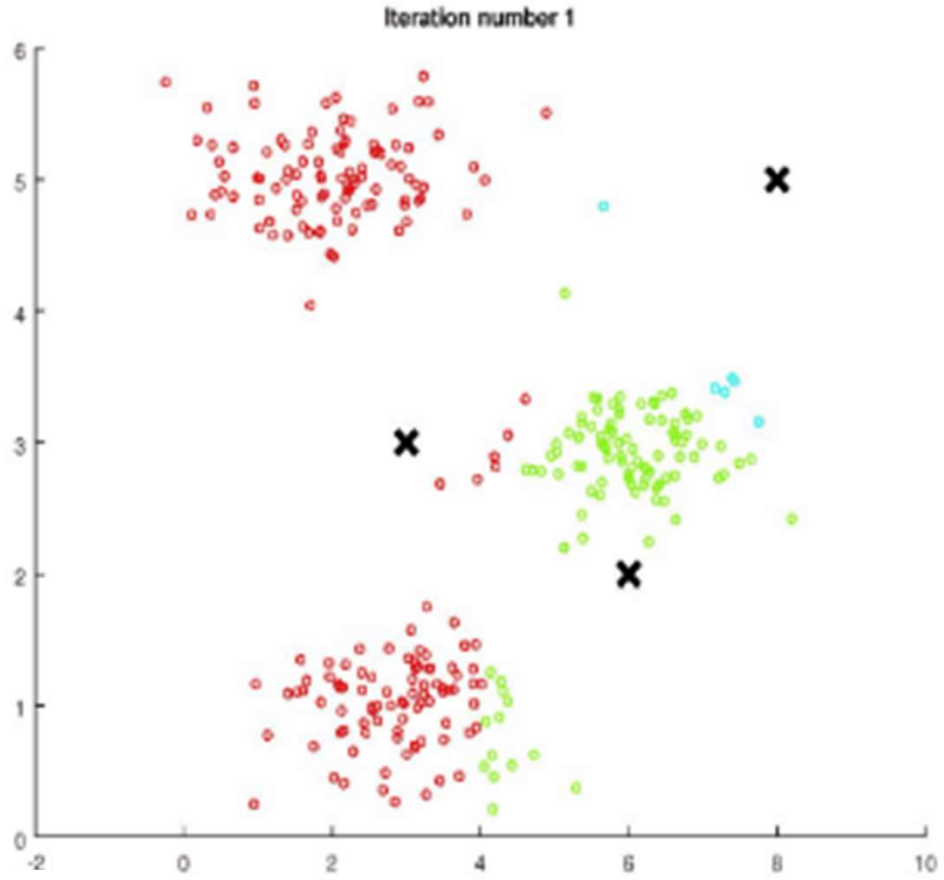


How it works – Step 4: Repeat steps 2 and 3

- Repeat steps 2 and 3 until no improvements are possible.



How k-Means Work




Thank
you !!!