

EE1390 INTRODUCTION TO AI and ML

K-Mean Clustering Algorithm

AYUSH KUMAR
(ES17BTECH11002)

February 2, 2020

Our goal is to Understand and Implement the K-mean clustering algorithm

K-Means Clustering

- *Attempts to split data into K groups that are closest to K centroid*
- * Unsupervised learning algorithm - uses only the positions of each data point*

K-Mean Clustering

It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong to that cluster) is at the minimum

K-Mean Clustering

The less variation we have within clusters, the more homogeneous (similar) the data points are within the same cluster.

The way K-means algorithm works is as follows:

step 1: Randomly pick K - clusters

step 2: Assign each data point to centroid its closest to

step 3: Recompute the centroid based on the average position of each centroids points

step 4: Iterate until points stop changing assignment to centroids.

Objective Function Used:

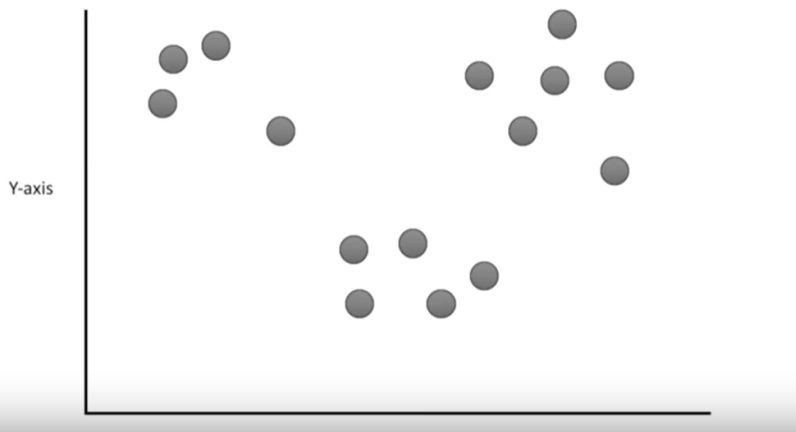
$$\mathbf{A} = \sum_{j=1}^k \sum_{i=1}^n ||\mathbf{x}_i^j - \mathbf{c}_j||^2$$

where $(\mathbf{x}_i^j - \mathbf{c}_j)^2$,

is a chosen distance measure between a data point and the cluster centre.

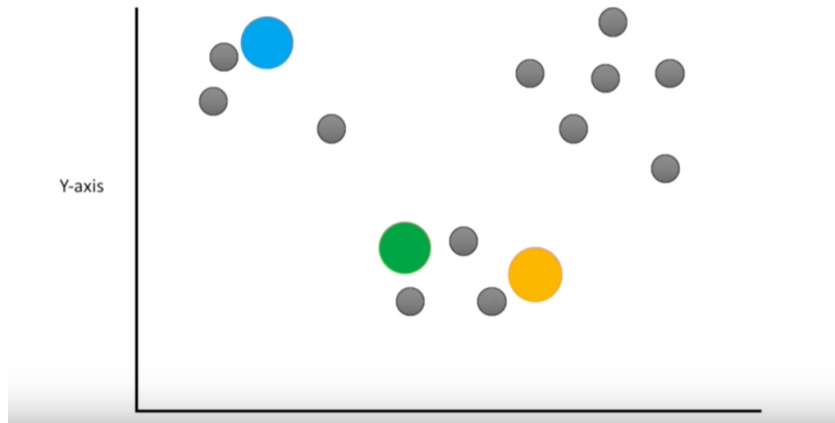
K-Mean Clustering

Sample point in space



K-Mean Clustering

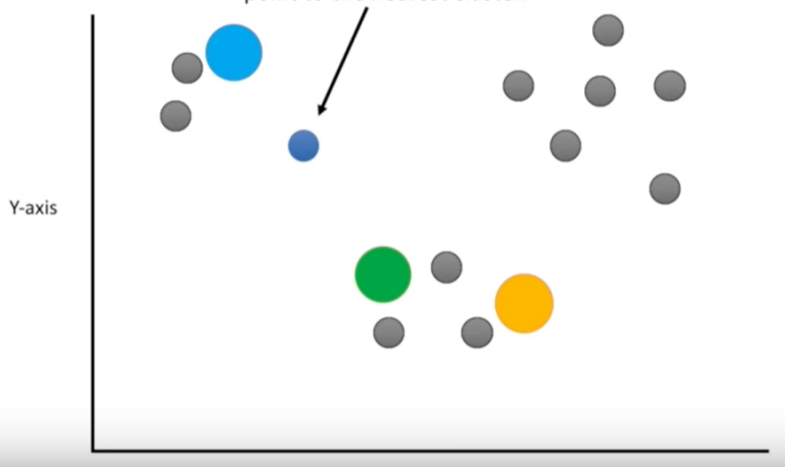
Pick three random point



K-Mean Clustering

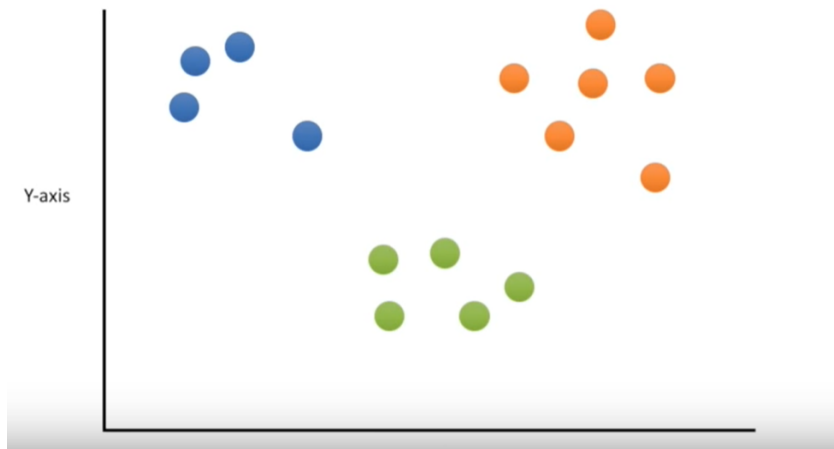
Assigning the point to nearest Cluster

point to the nearest cluster.



K-Mean Clustering

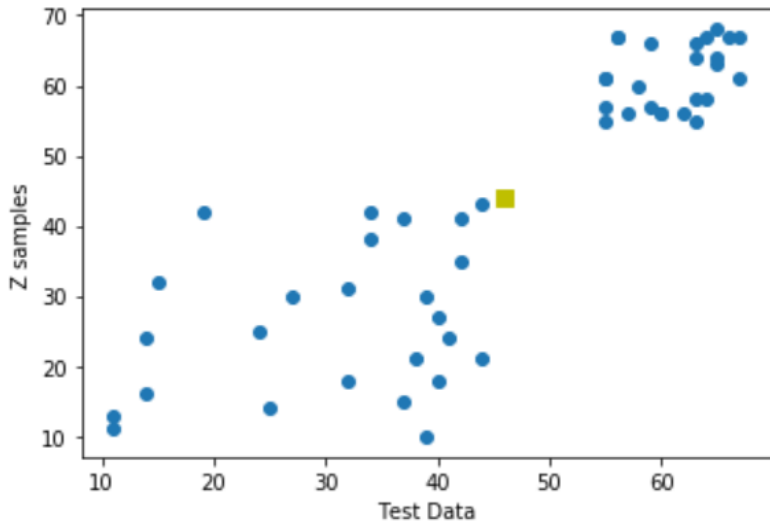
Cluster as per center and reclusters



K-Mean Clustering Example

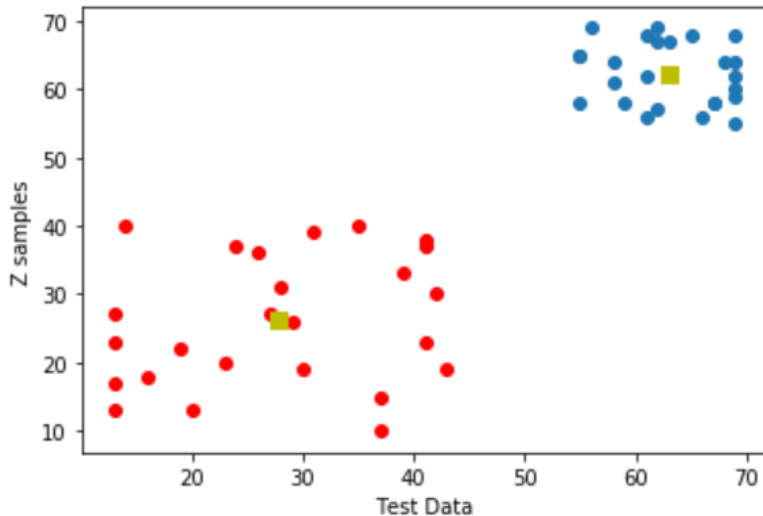
Pick random number from 1-45 and 55-70

Choosing $k = 1$



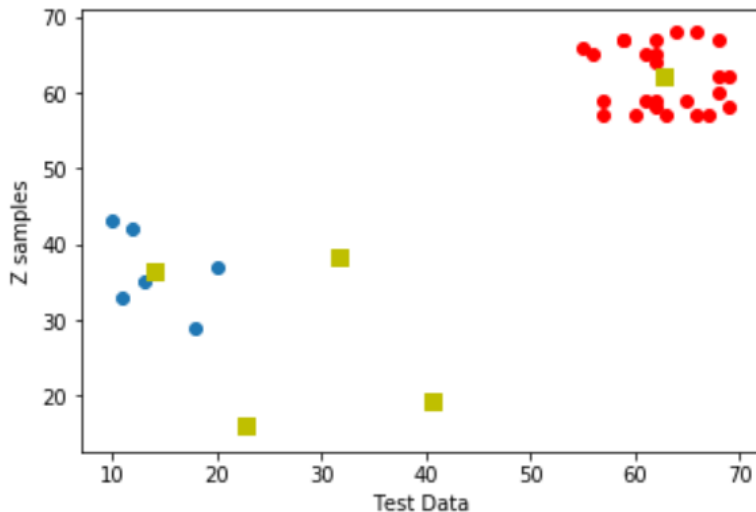
K-Mean Clustering Example

Choosing $k = 2$



K-Mean Clustering Example

Choosing $k = 5$



K-Mean Clustering Example

On Picking random number from 1-45 and 55-70

Conclusion :the best fit is on choosing $k=2$

K-Mean Clustering

Code link:

[*https://github.com/ayushkesh/k-mean-Clustering.git*](https://github.com/ayushkesh/k-mean-Clustering.git)