

Name: Rohit Kalyan

ID: 1002070724

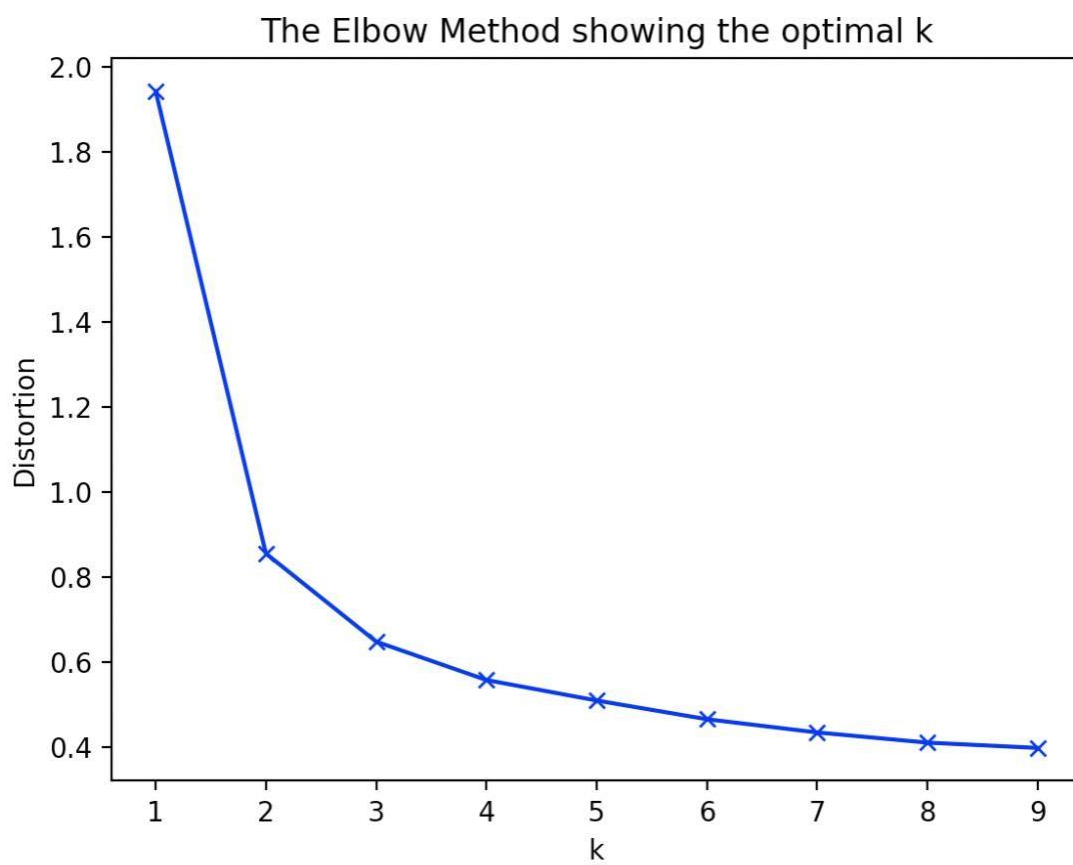
Assignment 3

- K means clustering makes K clusters from the data by updating the centroids in each iteration in order to make a perfectly clustered model. We use mean and standard deviations on the clusters in order to make a perfect model.
- We are using Iris data which has 5 columns in it which goes as Sepal Length, Sepal Width, Petal Length, Petal Width, category. We are making this into clusters in by processing the data with the K-means algorithm.
- We first use the Kmeans package in order to determine the optimum k value from the elbow graph. We got it to be 3.
- We now use this k=3 value and start clustering using the centroids we chose randomly in the beginning.
- We update the clusters by normalizing the distance between each centroid and data points and assign it to the closest centroid.
- After every iteration we update the centroids based and we continue the process till the centroids converge.
- The 3 final centroids of the 3 clusters are displayed in the output. Also Kmeans graph is plotted with the centroids pointed with black pyramids.

Results:

Elbow curve

Figure 1



KMeans graph(k=3):

