**K-means algorithm**

K-means is a unsupervised machine learning algorithm that groups similar pieces of information together (known as clustering) in order to find out underlying patterns. ‘K’ is defined as the fixed number of clusters within the data set and ‘means’ is defined as the data averaging.

K means is a popular algorithm due to its simplicity, it is used in market segmentation,

It works by first selecting random centroids which will be the start of all clusters, it then iterates through calculations in order to split the dataset into clusters and until there is no change in the centroids’ values. Computer the total squared distance between data points and centroids, give each data point to the closest centroid and take the average all the data points that belongs to each cluster.

image compression, call detail record (CDR) analysis and many more. In CDR, information such as the call, sms and internet activity of the user is used from telecom companies to find out the customers’ phone usage habits.