K-Means

This is an unsupervised learning algorithm which is used for clustering. For our case, we will be grouping the data in two classes with labels 1 and 0.

For each cluster, we will be defining a cluster centroid, which is basically the mean of the data belonging to that cluster.

is the centroid of the class ‘i’.

is the index of the cluster to which sample is assigned to.

First, we will initialize the centroids randomly or we can also randomly choose points from the training set. Then in each iteration, we will first assign each point in the training set to one of the classes based on the nearest centroid to that point. The distance used here is the Euclidean distance. Then we will update the centroids of each class to be the average of all the points assigned to that class.

, m = No. of samples

What we are actually doing here is, minimizing the cost function,

This is the total distance between the points in the training set and their corresponding cluster centroids.

In the first step in each iteration, we try to minimize this cost function by changing , i.e, by assigning the points to the cluster with the nearest centroid. Then in the next step we again try to minimize this cost by changing , i.e, by updating the centroids to be the mean of the points assigned to that cluster.

Then during prediction, we again assign the sample to the cluster whose centroid is closest to the sample.

One of the problems with KMeans is that with different initializations of initial centroids, we may end up with different clusters. To tackle this, during training we will run the algorithm more than once and then choose the set of centroids which gave us the minimum cost J.

The problem faced when using this on a labelled dataset is that the algorithm just classifies the training set into two groups. The actual labels and the labels assigned by the algorithm may be reversed. So in my code, I checked the accuracy of the model which is the fraction of prediction that is true and if it turns out to be less than 1, then I print the accuracy to be (1-accuracy). The predictions can also be returned as (1-predicion) in our case.