In k-means clustering, first, we have to randomly select a centroid for each class. Then, we iterate it until optimizing the position of the centroid for each class. We use Euclidean distance to find the centroid. Each data point from the training dataset is defined by taking Euclidean distance from the current centroid of each class and whichever centroid has the shortest distance from that data point. The new centroid is calculated by taking an average of the old centroid and data point. We define our stopping criteria as either centroid stop changing or a maximum number of iteration is reached.