Brian Griffin

K-Means Algorithm

The K-means algorithm’s objective is to group the data points of a sample into clusters that minimize the squared distances from the data’s points to their cluster’s center, also known as centroids. To achieve this, the algorithm must iterate through some steps. First, the algorithm must choose the number of centroids and place them in the data set. Instead of randomizing during the initialization process, it is generally best to spread them out to get the most defined clusters. After, the data points are grouped by their nearest centroid. Then, each centroid repositions itself so it is the average of all the data points in its group. These steps keep repeating until the centroids barely move from their previous position.