

ECE657 Tools of Intelligent Systems Design

Assignment 4 Question 2

Huy Trinh

July 2024

Question 3: K-means Limitations

Q3.1

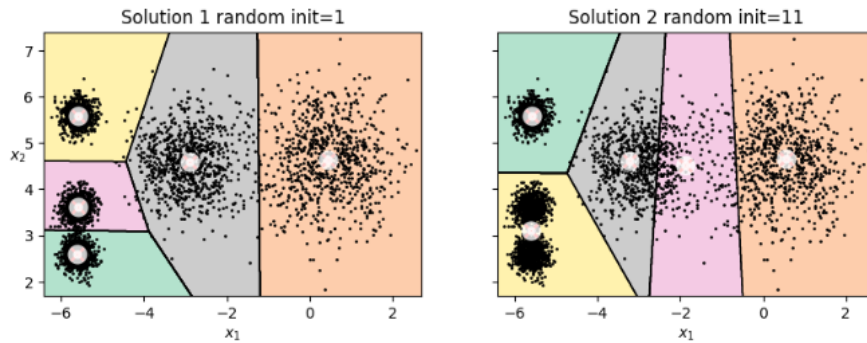


Figure 1: K-means algorithm using 2 different random states

Result Analytic and Potential Issue: The initial centroids chosen by the algorithm using *random_state* = 1 and *random_state* = 11 are different results in the final convergence of the algorithm. The issue is that random initialization of centroids can lead to inconsistent clustering result in different runs, which might not be best representation of the data structure. Another potential issue is convergence to local optima, which means the algorithm may not find the global optimum but suboptimal clustering.

Q3.2:

An improved method we propose here is **kmeans++**. **kmeans++** idea is choose initial points which are as far apart from each other as possible. The

method. Another method can be tried is foggy initialization which involves distributing the centroids more evenly across high-density areas.

Q3.3:

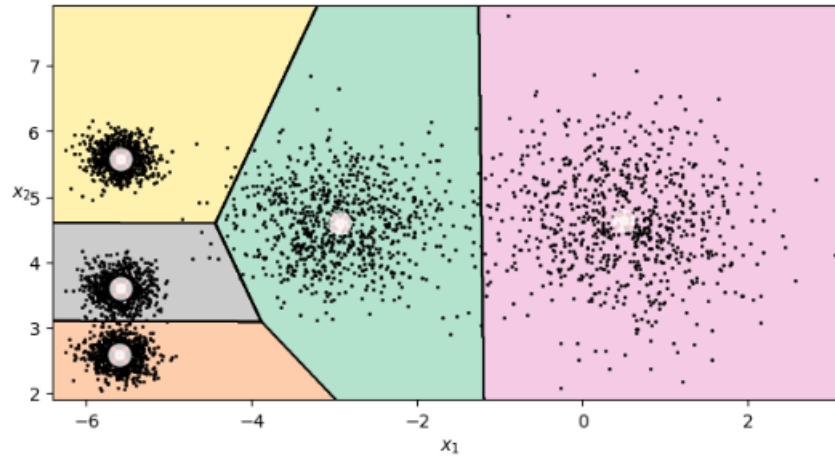


Figure 2: K-means algorithm using kmeans++ initialization

Using K-means++ initialization the algorithm converges to a better clustering solution with better positioned and more accurate compared to random initialization as above Q3.1.

Q3.4:

Some solutions to avoid in question is using **K-means++** initialization as recommended method to generate initial points for k-Means Algorithm. Other solutions are running K-means initialization multiple-times and choose the best result with lowest within-cluster sum of squares (WCSS) in elbow method given in the lecture. In Elbow method, we are varying number of clusters K from 1-10, for each K we calculate WCSS and plot it. When we analyze the graph, we can see that the graph will rapidly change at a point and thus creating an elbow shape. Another solution is we can use Foggy initialization, which is less common. In Foggy initialization, this attempt to avoid placing initial centroids in high-density areas while ensuring they are not too close to each other.