Machine Learning

CSCI 567 Spring 2021

Discussion: Clustering, K-means

Date: Mar 19th, 2021

1. Can k-means ever give results which contain more or less than k clusters? Given an example if it is possible.

2. Consider clustering n points on a line situated at coordinates equaling the powers of two $(1, 2, 4, 8, \ldots)$. Each point in a cluster of its own. Suppose at each step of the clustering algorithm we merge the two closest centroids. What is the centroid of the largest cluster after i steps. What is the total cost of clustering after i steps?