

An Example of Text Clustering with k -Means

Dell Zhang
24/03/2011

Given a set of five points $\{1, 2, 3, 4, 5\}$ on an axis, what clusters will the k -means clustering algorithm (with $k=2$) produce, if 4 and 5 are chosen as the two initial seeds?

Please use Euclidean distance. For example, the distance between points 2 and 5 is 3.



// *k*-means (*k*=2)

Step (1)

centres: 4 and 5
clusters: {1, 2, 3, 4} and {5}

Step (2)

centres: 2.5 and 5
clusters: {1, 2, 3} and {4, 5}

Step (3)

centres: 2 and 4.5
clusters: {1, 2, 3} and {4, 5}

Done!