Assignment 3 Report - KNN and Decision Trees

Helena Bales and Natalie Suderman

CS343 - Spring 2017

May 21, 2017

Contents

1 Non-Hierachical Clustering Implementing K-means algorithm

1.1 Implementing K-means with K of 2

The results of typical run of the kmeans algorithm when k is 2. I plotted multiple runs to show the trend.

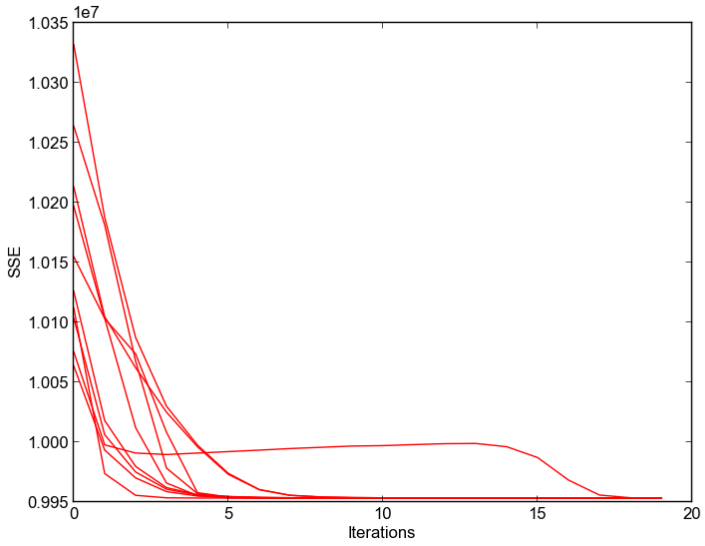

Figure 1: A typical run. SSE of Kmeans with k of 2 over number of iterations.

1.2 Apply K means to different values of K

The algorithm was tested by using different values of K (3, 4, 10, etc) and the minimum SSE from 10 runs for each K was plotted to demonstrate a trend in the size of the SSE vs number of K.

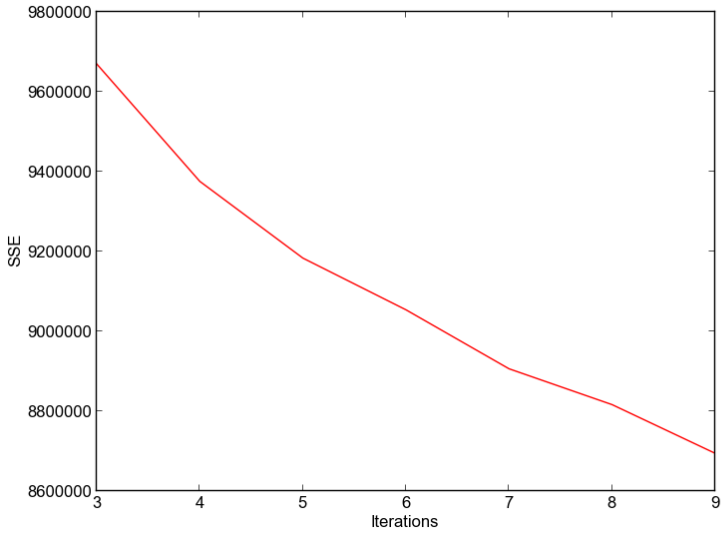

Figure 2: Min SSE found over 10 runs of the algorithm for each k tested.

2 Hierarchical Agglomerative Clustering

2.1 Compute HAC Using Single Link

2.2 Compute HAC Using Complete Link