EECS-HW1 Clustering, Classification, and Spark MLlib

Jing Qian (jq2282@columbia.edu)

October 3, 2019

Q1. Implement iterative K-means in Spark

(1) Within-cluster cost for each iteration using L1 distance.