Abstract

Although k-means works well in many cases it offers no accuracy guarantee and it has no idea to select ideal cluster representatives. This article presents a technique in which the initial cluster representatives in the standard k-means algorithm are chosen intelligently. Comparison of the quality of the clusters produced by the standard k-means algorithm, k-means using Furthest-First, and k-means using the proposed initialization technique have investigated. Experiment result shows that the quality of the clusters improves with the proposed algorithm in most of the cases.
- Shuttle Dataset Available: http://mlr.cs.umass.edu/ml/datasets/stalog+(shuttle)

Index Terms

Computer Science
Algorithms

Keywords
Cluster representative  cluster quality  Furthest-First Technique  centroid