Hybrid Imperialist Competitive Algorithm and Dynamic Validity Index to find the best clusters

Mojgan Ghanavati[1]

mojgan.ghanavati@gmail.com

Behrouz Minaei[3]

b_minaei@iust.ac.ir

Mohammad Reza Gholamian[2]

gholamian@iust.ac.ir

Mehran Davoudi[4]

mehran.davoudi@gmail.com

Photograph of Presenter

[1,2,3,4] Iran University of science and technology
Paper Reference Number: 16
Name of the Presenter: Mehran Davoudi

Abstract

Cluster analysis is one of attractive data mining technique that use in many fields. One of the popular types of clustering algorithms is the center based clustering algorithm. K-means used as a popular clustering method due to its simplicity and high speed in clustering large datasets. However, K-means has two shortcomings. K-means is dependent on the initial state and convergence to local optima in some of the large problems. In order to these shortcomings, in an unsupervised clustering the number of clusters needs to be fixed by a human analyst too. In order to overcome local optima problem and for determining the number of clusters, lots of studies done in clustering. In this paper we combine a new search heuristic called "Imperialist Competitive Algorithm" with "Dynamic Validity Index (DVIndex)" to find the best clusters. In this algorithm, we assume each clustering solution as a country and use we use DVIndex as an efficient method to find number of clusters for calculating the clustering cost in each step. We compared proposed algorithm with other heuristics algorithm in clustering, such as traditional K-means, CSO, GKA and PSO-GA, by implementing them on several well-known datasets. Our findings show that the proposed algorithm works better than the others.

Key words: Clustering, Meta-heuristic, K-means, Imperialist Competitive, DVIndex

1. Introduction