



Speedy Algorithm for Clustering Imbalanced Data

M. H. Marghny, Rasha M. Abd El-Aziz

Abstract:

Fast Balanced K-means (FBK-means) clustering approach is one of the most important consideration when one want to solve clustering problem of balanced data. Mostly, numerical experiments show that FBK-means is faster and more accurate than the K-means algorithm, Genetic Algorithm, and Bee algorithm. FBK-means Algorithm needs few distance calculations and fewer computational time while keeping the same clustering results. However, the FBK-means algorithm doesn't give good results with imbalanced data. To resolve this shortage, a more efficient clustering algorithm, namely Fast K-means (FK-means), developed in this paper. This algorithm not only give the best results as in the FBK-means approach but also needs lower computational time in case of imbalance data.

Keywords:

Clustering, K-means Algorithm, Bee Algorithm, Genetic Algorithm, FBK-means Algorithm, FK-means Algorithm

Published In:

CiiT International Journal of Data Mining and Knowledge Engineering , 7-2 , 82-88