

EFFICIENT CLUSTERING OF DATASET BASED ON PARTICLE SWARM OPTIMIZATION

SURESH CHANDRA SATAPATHY¹ & ANIMA NAIK²

¹Sr. Member IEEE, ANITS, Vishakapatnam, Andhra Pradesh, India

²Majhighariani Institute of Technology and Sciences, Rayagada, Odisha India

ABSTRACT

The Automatic Particle Swarm Optimization (AUTO-PSO) clustering algorithm can generate more compact clustering results than the traditional K-means clustering algorithm. However, when clustering high dimensional datasets, the AUTO-PSO clustering algorithm is notoriously slow because its computation cost increases exponentially with the size of the dataset dimension. Dimensionality reduction techniques offer solutions that both significantly improve the computation time, and yield reasonably accurate clustering results in high dimensional data analysis. In this paper, we present an idea of an algorithm that can combine dimensionality reduction techniques of weighted PCs with Auto PSO for clustering. The intention is to reduce complexity of datasets and speed up the Auto PSO clustering process. We report significant improvements in total runtime. Moreover, the clustering accuracy of the dimensionality reduction AUTO-PSO clustering algorithm is comparable to the one that uses full dimension space.

KEYWORDS: Clustering, Particle Swarm Optimization, Principal Component, Dimension Reduction