

## Abstrak

Klasterisasi (*clustering*) merupakan salah satu fungsionalitas data mining yang digunakan untuk melakukan pengelompokan data ke dalam suatu kelas atau *cluster*. Prinsip dasar klasterisasi adalah mengelompokkan objek pada suatu kelas yang memiliki kemiripan sangat besar dengan objek lain pada kelas yang sama (*similarity*), tetapi sangat tidak mirip dengan objek pada kelas lain (*dissimilarity*). Terdapat beberapa teknik klasterisasi antara lain: metode Partisi (*K-means Clustering*), metode Hierarki (*Divisive and Agglomerative Clustering*), metode *Density-Based (DBSCAN)*, dan sebagainya.

Pada tugas akhir ini, digunakan metode Partisi dengan algoritma PSO + K-means (*Particle Swarm Optimization + K-means*) yang merupakan penggabungan antara algoritma K-means dan algoritma PSO (*particle Swarm Optimization*). Metode pengklasteran dengan menggunakan K-means sangat banyak digunakan untuk mengelompokkan data dengan similaritas yang tinggi. Akan tetapi K-means memiliki kelemahan dalam penentuan titik pusat inisial klaster yang dilakukan secara random sehingga sering kali menyebabkan terjebak pada lokal optimal dan hasil pengklasteran menjadi tidak optimal. Untuk lebih mengoptimalkan penentuan titik pusat maka digunakanlah algoritma PSO + K-means. PSO + K-means merupakan algoritma yang diusulkan oleh Xiaohui Cui pada tahun 2005. Algoritma ini selalu menghasilkan konvergensi pada global optimal. PSO + K-means mampu menghindari lokal optimal. Dalam tugas akhir ini juga akan dilakukan perbandingan evaluasi hasil klasterisasi yang dihasilkan oleh klasterisasi menggunakan metode K-means dan PSO dengan menggunakan ADVDC (*average distance documens cluster centroid*) dan *time performance*.

**Kata kunci** : *clustering, particle Swarm Optimization, PSO, K-means, similarity.*