

## 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 FGKA (*Fast Genetic K-means Algorithm*) yang merupakan penggabungan antara algoritma K-means dan algoritma Genetika. 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 dan dengan waktu yang seminimal mungkin maka digunakanlah algoritma FGKA (*Fast Genetic K-means Algorithms*). FGKA merupakan algoritma yang dikembangkan dari algoritma GKA (*Genetic K-means Algorithm*) yang diusulkan oleh Yi Lu pada tahun 2004. Algoritma ini selalu menghasilkan konvergensi pada global optimal. FGKA dan GKA mampu menghindari lokal optimal akan tetapi FGKA berjalan lebih cepat dibandingkan GKA. Dalam tugas akhir ini juga akan dilakukan perbandingan evaluasi hasil klasterisasi yang dihasilkan oleh perangkat lunak klasterisasi menggunakan metode K-means.

**Kata kunci** : *clustering, GKA, K-means, FGKA, Fast Genetic K-means Algorithm*