

PROGRAM APLIKASI BAHASA R UNTUK PENGELOMPOKAN OBJEK MENGGUNAKAN METODE *K-MEDOIDS CLUSTERING*

ABSTRAK

Cluster atau dalam bahasa Indonesia dapat ditulis sebagai klaster adalah nama lain dari kelompok. Pada hakekatnya, analisis klaster adalah salah satu alat pengelompokan (*clustering*) dalam analisis statistika multivariat yang pada akhirnya menghasilkan sejumlah klaster (kelompok). Pengelompokan dilakukan pada objek/pengamatan (baris) dalam data yang memiliki kemiripan sangat besar dengan objek/pengamatan lainnya dalam satu klaster. Kemiripan tersebut diukur menggunakan ukuran kemiripan. Analisis klaster terbagi menjadi metode hierarki dan metode non-hierarki. Metode hierarki menggunakan struktur seperti pohon pada keputusan pengelompokan objeknya. Metode non-hierarki menempatkan objek-objek ke pusat klaster terdekat ketika banyak klaster sudah ditentukan terlebih dahulu. Ada dua buah asumsi yang harus dipenuhi agar analisis klaster dapat digunakan, yaitu data bebas pencilan (*outlier*) dan bebas dari masalah multikolinearitas. Namun, pada kenyataannya asumsi pencilan sering kali tidak terpenuhi. Salah satu metode analisis klaster yang tahan terhadap pencilan adalah metode *k-medoids* atau biasa dikenal dengan nama *Partitioning Around Medoids* (PAM). *K-medoids* merupakan salah satu bagian dari metode non-hierarki. Untuk menyelesaikan pengelompokan objek menggunakan *k-medoids*, diperlukan iterasi yang tidak sedikit. Jika menggunakan cara manual, tentu memakan banyak waktu dan kemungkinan mengalami kesalahan yang cukup tinggi. Untuk mengatasi hal tersebut, penelitian menggunakan bahasa pemrograman R yang dirancang untuk membentuk sebuah program aplikasi pengelompokan objek berdasarkan metode *k-medoids* pada suatu data. Hasil dari penelitian ini adalah sebuah program aplikasi dalam bahasa R. Program ini dapat digunakan sebagai alat bantu pengelompokan objek dengan memperlihatkan proses iterasi dari pengelompokan sampai akhirnya terbentuk sejumlah klaster yang sebelumnya telah ditentukan.

Kata Kunci: Analisis Klaster, Pengelompokan, *K-Medoids*, PAM, bahasa R.

R LANGUAGE APPLICATION PROGRAM FOR OBJECT CLUSTERING USING *K-MEDOIDS CLUSTERING*

ABSTRACT

Cluster or in Indonesian can be written as a 'klaster' is another name of the group. In essence, cluster analysis is one of the clustering tools in multivariate statistical analysis that ultimately produces a number of clusters (groups). Clustering is done on objects / observations (rows) in data that have very similar resemblance to other objects / observations in a cluster. The similarities are measured using similarity measurement. Cluster analysis is divided into hierarchical methods and non-hierarchical methods. The hierarchical method uses a tree-like structure in its object clustering decisions. The non-hierarchical method places objects to the nearest cluster center when the number of clusters are predefined. There are two assumptions that must be fulfilled for cluster analysis to be used, those are the data free from outliers and free from multicollinearity problems. In reality, however, the assumption of outliers is often unfulfilled. One method of cluster analysis that is robust to outliers is k-medoids method or commonly known as Partitioning Around Medoids (PAM). K-medoids are one part of a non-hierarchical method. To accomplish the object grouping using k-medoids, it takes a lot of iteration. If using a manual way, would take a lot of time and the possibility of experiencing a fairly high error. To overcome this, research using R programming language designed to form an object clustering application program based on k-medoids method on a data. The result of this research is an application program in R language. This program can be used as a tool clustering object by showing the iteration process from clustering until finally formed a number of clusters that have been previously determined.

Keywords: *Cluster Analysis, Clustering, K-Medoids, PAM, R.*