

ABSTRAK

Algoritma PAM (*Partitioning Around Medoids*)-*Lite* dengan Program R untuk Data Berukuran Besar

Oleh

Rizki Agung Wibowo

Analisis klaster k-medoid atau dikenal sebagai PAM (*Partitioning Around Medoid*) adalah salah satu analisis klaster berbasis partisi. PAM memiliki keunggulan yaitu *robust* terhadap data pencilan akan tetapi memiliki kelemahan yaitu tingkat komputasi yang buruk pada data berukuran besar, sehingga dikembangkan algoritma PAM-*Lite* untuk menjadi alternatif PAM. Pada penelitian ini dilakukan konstruksi *function* PAM-*Lite* pada program R karena belum ditemukan *package* yang didalamnya terdapat *function* PAM-*Lite* kemudian menguji keefektifannya menggunakan data simulasi. Berdasarkan hasil yang didapat, *function* PAM-*Lite* yang telah dikonstruksi dapat diproses dengan baik pada program R dan berdasarkan R-*square*, lebar *silhouette* dan waktu proses, PAM-*Lite* lebih efisien digunakan jika dibandingkan dengan PAM pada data berukuran besar.

Kata kunci : Data pencilan, Analisis klaster, PAM-*Lite*, Pemrograman R

ABSTRACT

PAM (*Partitioning Around Medoids*)-*Lite* algorithm with R for a Massive Data

By

Rizki Agung Wibowo

K-Medoid cluster analysis or known as PAM (Partitioning Around Medoid) is a partition-based cluster analysis method. This method has the advantage namely robust against outliers, however it also has disadvantage i.e. it is worse in computation when applied on large data, for this reason the PAM-Lite algorithm was developed to be an alternative for clustering large data. In this study, we constructed an R function (i.e. a block of codes which only runs when it is called) of PAM-Lite algorithm considering the unavailability of R packages that contains PAM-Lite function. The effectiveness of the constructed function then was tested using basic cluster benchmark data. Based on the result, PAM-Lite function can process properly in R program and is more efficient than PAM based on the R Square, Silhouette width and processing time on large data.

Keyword : Outlier, Cluster Analysis, PAM-Lite, R Programming