

ABSTRAK

Analisis kluster adalah salah satu metode interdependensi dalam analisis statistik multivariat yang mengelompokkan observasi ke dalam beberapa kluster. Susunan kluster yang baik akan mempunyai homogenitas yang tinggi antar anggota dalam satu kluster dan heterogenitas yang tinggi antar kluster. Terdapat dua metode pengelompokan dalam analisis kluster, yaitu metode hirarki dan non-hirarki. *K-Means* merupakan metode non-hirarki yang cukup terkenal dan paling banyak digunakan, namun karena menggunakan rata-rata sebagai *centroid* atau titik pusat klasternya, metode ini lebih sensitif terhadap keberadaan pencilan pada data. Di sisi lain, median adalah statistik deskriptif yang cenderung lebih tahan terhadap pencilan, sehingga berkembanglah metode yang dapat mengelompokkan data yang mengandung pencilan yaitu metode *K-Medians*, yang menggunakan median sebagai *centroid*. Sebagai contoh penerapan metode *K-Medians* digunakan pada data jumlah tenaga kesehatan menurut kabupaten/kota di Provinsi Jawa Barat tahun 2015 yang menghasilkan 5 kluster optimal dengan masing-masing karakteristik klasternya yang berbeda.

Kata kunci: analisis kluster, ukuran kemiripan, *K-Medians Clustering*.

ABSTRACT

Cluster analysis is an interdependence multivariate statistic method that classify observations into several clusters. Good clusters will have a high homogeneity among members in a cluster and high heterogeneity between clusters. There are two methods of grouping in the cluster analysis, hierarchical and non-hierarchical clustering. K-Means clustering method is quite well known and most widely used, but because it uses the average as centroid, this method is more sensitive to outliers in the data. In the other hand, median is a descriptive statistic which more resistant to outliers, and K-Medians clustering is the solution. Use a median as the centroid, this method owes its use to robustness of the median as a statistic. As an example, application of this method used in the data on the number of health workers by district/city in Jawa Barat Province in 2015 that resulted five clusters optimal, with each clusters have different characteristics.

Keywords: cluster analysis, similarity, K-Medians clustering.