

Junaedi Asmoro, 2005. *Pengelompokan Data Kategorik Dengan Algoritma K-Modus*. Skripsi ini dibawah bimbingan Ir. Dyah Herawatie, M.Si dan Drs. Eto Wuryanto, DEA. Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Airlangga Surabaya.

---

## ABSTRAK

Skripsi ini bertujuan untuk melakukan pengelompokan data berjenis kategori dengan menggunakan metode non-Hirarki berdasarkan algoritma K-Modus dalam analisis kelompok. Algoritma K-Modus merupakan pengembangan dari algoritma K-Mean dengan mengganti mean dengan modus sebagai pusat kelompoknya.

Pengelompokan data dengan algoritma K-Modus dilakukan secara iterasi dengan empat langkah dasar. Langkah yang pertama yaitu menentukan pusat kelompok (modus) awal. Langkah kedua, menghitung jarak antar obyek pengamatan dengan pusat kelompok. Langkah ketiga yaitu menempatkan obyek-obyek pengamatan pada kelompok yang mempunyai jarak terdekat (minimal) dengan pusat kelompoknya. Langkah yang keempat menentukan modus yang baru dari masing-masing kelompok, proses selanjutnya kembali ke langkah kedua. Kemudian proses ini diulang sampai tidak ada obyek yang berpindah ke kelompok lain.

Program S-plus yang telah dibuat diterapkan pada dua data sekunder. Pada pengelompokan data I baik menggunakan metode I dan II didapatkan hasil pengelompokan yang sama jika nilai  $k = 2$ , sedangkan pada pengelompokan data II didapatkan hasil pengelompokan yang berbeda jika nilai  $k = 3$ . Perbedaan nilai awal dalam menetapkan pusat kelompok dapat memberikan hasil yang berbeda dalam proses pengelompokan

Kata kunci : K-Mean, K-Modus, Data kategorik, Analisis kelompok

Junaedi Asmoro, 2006. Clustering of Categorical Data by using K-Modes Algorithm. The *skripsi* was under guidance of Ir. Dyah Herawatie, M.Si. dan Drs. Eto Wuryanto, DEA. Mathematics Department. Faculty of Mathematics and Natural Science. Airlangga University.

---

## ABSTRACT

The purpose of this *skripsi* is to cluster the categorical data by using non-hierarchical method based on K-Modes algorithm in cluster analysis. K-Modes algorithm is the improvement of K-Means algorithm by changing means with modes as the center of cluster.

Clustering of data by using K-Modes algorithm can be applied by iteration with four basic steps. The first step is to determine the first center of cluster (modes). Determination of the first center of cluster can be done by using two kind of methods. The second step is to calculate the distance between observation object with the center of cluster. The third step is to allocate observation objects in the cluster which has nearest distance (minimum) with the center of cluster. The fourth step is to determine the new modes from each cluster, the next process is back to the second step. These process is repeated until there is no object that move to another cluster.

The S-Plus program which have been made applied at two secondary data. In the clustering of data I by using method I or II was obtained the same result of clustering if the value of  $k$  is 2, whereas in clustering of data II was obtained the different result of clustering if the value of  $k$  is 3. The different of first value in determining the center of clustering can give different result in process of clustering.

Key words : K-Mean, K-Modes, Categorical data, Cluster analysis