

Kurniawan Setya B, 2006. Penggunaan Algoritma *CLOPE* Pada Pengelompokan Data Kategorik. Skripsi ini dibawah bimbingan Ir. Dyah Herawatie, M.Si dan Drs. Eto Wuryanto, DEA. Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Airlangga.

ABSTRAK

Skripsi ini bertujuan untuk melakukan pengelompokan data berjenis kategorik dengan menggunakan algoritma *CLOPE* (*Clustering with sLOPE*). Algoritma *CLOPE* khusus digunakan untuk mengelompokkan data transaksional atau data biner yang mempunyai dimensi dan volume yang besar (data mining). Algoritma ini didasarkan pada ide peningkatan rasio tinggi-lebar (*slope*) dari histogram kelompok, kemudian digeneralisasi dengan parameter *repulsion* yang mengontrol kehomogenitasan transaksi-transaksi dalam sebuah kelompok.

Pengelompokan data kategorik dengan algoritma *CLOPE* dimulai dengan tahap peletakan yaitu dengan membentuk pengelompokan awal, yang diperoleh melalui *Profit* (fungsi ukur) yang maksimum. Tahap selanjutnya, yaitu tahap pemindahan. Pada tahap ini, masing-masing obyek pengamatan dipindahkan kedalam kelompok yang telah ada pada pengelompokan awal. Tahap pemindahan diperoleh dengan mengoptimalkan *Profit* (fungsi ukur). Jika tidak ada perubahan pengelompokan dari iterasi sebelumnya maka algoritma berhenti dengan pengelompokan terakhir sebagai output.

Dari program *S-Plus* yang telah dibuat, diterapkan pada dua data sekunder. Pada data I dengan menggunakan algoritma *CLOPE*, pengelompokan yang terbentuk sebanyak dua kelompok dengan *repulsion* $r = 2$. Sedangkan dari pengelompokan 50 dan 100 obyek data II menunjukkan bahwa waktu penyelesaian program dari algoritma *CLOPE* lebih cepat (efisien) daripada algoritma *QROCK* dan jauh lebih cepat daripada algoritma *ROCK*.

Kata kunci : algoritma *CLOPE*, data kategorik, analisis kelompok, data mining

Kurniawan Setya B, 2006. The use of CLOPE algorithm for Categorical Data Clustering. This *skripsi* is under guidance of Ir. Dyah Herawatie, M.Si and Drs.Eto Wuryanto, DEA. Mathematics Department Faculty of Mathematics and Natural Science. Airlangga University.

ABSTRACT

The purpose of this *skripsi* is to cluster categorical data by using CLOPE (Clustering with sLOPE) algorithm. CLOPE algorithm is used especially for clustering transactional data or binary data that have dimensionality and large volume. This algorithm is based on the idea of increasing the height-to-width ratio (slope) of the cluster histogram, then it is generalized by a repulsion parameter that controls tightness of transactions in a cluster.

Clustering of categorical data by using CLOPE algorithm is started with “putting phase” that build the initial clustering, driven by maximum Profit (the criterion function). The next phase is “moving phase”. In this phase, each observation object is moved to an existing cluster in the initial clustering. “Moving phase” is obtained by optimizing the Profit (the criterion function). If there is no changes of the clustering in a previous iterations, hence the algorithm will stop and the final clustering will be the output.

The S-Plus program, that is implementation of CLOPE algorithm, is used to process two secondary data. The result of first data is two groups with *repulsion* $r = 2$. While from the clustering 50 and 100 object of second data shows that the computing time of the CLOPE algorithm is faster (efficient) than QROCK algorithm and it is extremely faster than ROCK algorithm.

Key word : CLOPE algorithm, categorical data, cluster analysis, data mining.