



PENENTUAN JUMLAH OPTIMAL KELOMPOK MELALUI BEBERAPA METODE PENGELOMPOKAN DAN INDEK KRITEIRA

Oleh :
Kariyam

Dibawah bimbingan :
Prof. Drs. Subanar, Ph.D

INTISARI

Pada artikel ini akan memperhatikan masalah penentuan jumlah optimal kelompok. Untuk menangani masalah tersebut, kami tertarik pada penggunaan statistik *gap*. Teknik yang digunakan menggunakan hasil beberapa metode pengelompokan, khususnya *K-Means*, hirarki dengan pautan tunggal dan metode Ward.

Keefektifan statistik *gap*, akan dibandingkan dengan statistik *silhouette* dan beberapa metode yang telah diusulkan oleh Calinski-Harabazs, Hartigan, dan Krzanowski-Lai. Perbandingan didasarkan pada model dengan variasi jumlah variabel dan ukuran sampel. Hasil dari simulasi menunjukkan bahwa statistik *gap* tidak jauh berbeda dengan metode-metode yang telah diusulkan.

Kata-kata kunci : metode pautan tunggal, metode Ward, metode *K-Means*, statistik *gap*, statistik *silhouette*





ESTIMATING THE OPTIMAL NUMBER OF GROUPS VIA SOME CLUSTERING METHODS AND CRITERION INDEXES

By :
Kariyam

Under Supervision :
Prof. Drs. Subanar, Ph.D

ABSTRACT

This article considers the problem of estimating the optimal number of groups. We interested in using the gap statistic for overcome it. The technique uses the output of some clustering methods, specially K-Means, hierarchical with single linkage and ward methods.

The effectiveness of the gap statistic is compared with silhouette statistic and many methods have been proposed for estimating the number of clusters by Calinski-Harabazs, Hartigan, and Krzanowski-Lai. Comparisons based on models with various the number of variables and sample sizes. The result of simulation shows that the gap statistic usually not distinct with other methods.

Key words : *single linkage method, Ward's method, K-Means method, the gap statistic, silhouette statistic*

