

ABSTRACT

Selection Simulation of The Best Hierarchical Agglomerative Clustering Method between Average Linkage and Ward's on Data Containing Multicollinearity Problems

By

Rizki Agung Wibowo

Multicollinearity is a linear relation (collinearity) that exists between independent variables. In cluster analysis, the effect is different because multicollinearity is a form of implicit weighting. Principal component analysis can be used to reduce the number of variables that correlated into the number of new variables that uncorrelated by maintaining as much variety of data, by using the result of principal component analysis, we can do cluster analysis by average linkage and Ward's methods, then the best method will be chosen based on Dunn and RS indices, it was concluded that Ward's method is better than average linkage based on RS index which means that cluster formed using Ward's has more different characteristics than average linkage method while using Dunn index it can be concluded that average linkage is better than Ward's method which means that cluster formed using average linkage has more compactness than Ward's method.

Keyword : Multicollinearity, Average Linkage Method, Ward's Method,
Principal Component Analysis, Dunn Index, RS Index

ABSTRAK

Simulasi Pemilihan Metode Analisis *Cluster* Hirarki *Agglomerative* Terbaik antara *Average Linkage* dan Ward pada Data yang Mengandung Masalah Multikolinearitas

Oleh

Rizki Agung Wibowo

Multikolinearitas adalah hubungan linear yang ada di antara variabel independen, pada analisis kluster efek yang ditimbulkan oleh multikolinearitas berbeda, dikarenakan pada dasarnya multikolinearitas adalah bentuk pembobotan implisit. Analisis komponen utama dapat digunakan untuk mereduksi jumlah himpunan peubah yang banyak dan saling berkorelasi menjadi peubah-peubah baru yang tidak berkorelasi dengan mempertahankan sebanyak mungkin keragaman data tersebut, dengan menggunakan hasil analisis komponen utama dilakukan analisis kluster menggunakan metode *average linkage* dan Ward, yang kemudian akan dipilih metode terbaiknya berdasarkan nilai indeks Dunn dan indeks R-Sq (RS), didapat kesimpulan bahwa metode Ward lebih baik dibandingkan *average linkage* yang ditinjau berdasarkan indeks RS yang berarti kluster yang terbentuk dengan menggunakan metode Ward memiliki karakteristik yang lebih berbeda dibanding dengan metode *average linkage*, sedangkan dengan menggunakan indeks Dunn

didapatkan kesimpulan bahwa metode *average linkage* lebih baik dibandingkan metode Ward yang berarti klaster yang terbentuk dengan menggunakan metode *average linkage* lebih kompak dibanding dengan metode Ward.

Kata kunci : Multikolinearitas, Metode *Average Linkage*, Metode Ward, Analisis Komponen Utama, Indeks Dunn, Indeks RS