

ABSTRACT

IMPLEMENTATION OF THE ENSEMBLE ROBUST CLUSTERING USING LINKS (ROCK) METHOD FOR SENIOR HIGH SCHOOL (SMA) CLUSTERIZATION IN BANDAR LAMPUNG

By

Dini Desita

Cluster analysis is a method in interdependent multivariate analysis to group a set of objects based on their similar characteristics. One of the cluster analysis methods that can be used to group mixed data types (numeric and categorical) is Ensemble ROCK which has the advantage that there is a better accuracy value compared to the agglomerative hierarchical method because it has good scalability and has better quality in handle categorical data. In this analysis, the K-Medoids method is used to form clusters on numerical data and the ROCK method is used on categorical data. The purpose of this study is to apply the Ensemble ROCK cluster analysis to classify high schools in Bandar Lampung and obtain the best number of clusters. The data used is data from 60 high schools in Bandar Lampung based on high school facilities and infrastructure as well as human resources. The best grouping results are selected based on the ratio value of S_w and S_b . From the results and analysis it was found that the number of clusters-2 with a value of $\theta=0.05$ was the best cluster result in mixed data with the smallest S_w and S_b ratio value of 0.0310. The characteristics of cluster-1 are clusters consisting of 25 schools which have better quality than cluster-2 which consists of 35 schools.

Keyword: Cluster Analysis, *Ensemble ROCK*, *K-Medoids*, Senior High School, Ratio S_w dan S_b

ABSTRAK

IMPLEMENTASI METODE *ENSEMBLE ROBUST CLUSTERING USING LINKS (ROCK)* UNTUK KLASTERISASI SEKOLAH MENENGAH ATAS (SMA) DI BANDAR LAMPUNG

Oleh

Dini Desita

Analisis *cluster* merupakan salah satu metode dalam analisis multivariat yang digunakan untuk mengelompokkan sekumpulan objek berdasarkan kemiripan karakteristiknya. Metode *ensemble ROCK* adalah salah satu teknik dalam analisis *cluster* yang dapat digunakan untuk mengelompokkan data dengan tipe campuran (numerik dan kategorik). Metode ini memiliki ketepatan dan sifat skalabilitas yang baik untuk data berskala campuran. Dalam penelitian ini, metode *K-Medoids* digunakan untuk membentuk *cluster* pada data numerik dan metode *ROCK* digunakan pada data kategorik. Tujuan dari penelitian ini adalah menerapkan analisis *cluster ensemble ROCK* untuk mengelompokkan SMA di Bandar Lampung berdasarkan sarana dan prasarana serta sumber daya manusia SMA dengan menggunakan metode *ensemble ROCK*. Data yang digunakan adalah data 60 SMA di Bandar Lampung. Hasil pengelompokkan terbaik dipilih berdasarkan nilai *ratio Sw dan Sb*. Dari hasil dan analisis data diperoleh bahwa pada nilai $\theta = 0.05$ dihasilkan nilai *ratio Sw dan Sb* terkecil sebesar 0.0310 sehingga dapat disimpulkan bahwa jumlah *cluster* optimum adalah 2. Dengan karakteristik *Cluster-1* merupakan *cluster* yang terdiri dari 25 sekolah yang memiliki kualitas yang lebih baik dibandingkan pada *Cluster-2* yang terdiri dari 35 sekolah.

Kata Kunci: Analisis *Cluster*, *Ensemble ROCK*, *K-Medoids*, Sekolah Menengah Atas, *Ratio Sw dan Sb*