

ABSTRAK

Clustering merupakan metode *Unsupervised Learning* yang digunakan untuk pengelompokan pada data yang belum diketahui labelnya. *K-means* merupakan salah satu metode *Clustering* yang memiliki kemampuan mengelompokan data relatif cepat, tapi memiliki kesulitan menentukan *Centroid* awalnya. Maka fungsi dari metode *Hierarchical* ini untuk membantu memecahkan masalah penentuan *Centroid* awal dari metode *K-means*. Penelitian ini bertujuan untuk menerapkan kombinasi *Hierarchical* dan *K-means* untuk mengelompokan nilai UASBN SD dan mengetahui seberapa bagus hasil evaluasi dari *Hierarchical K-means*. Data nilai UASBN SD yang dipergunakan memiliki atribut nilai Bahasa Indonesia, Matematika dan IPA. Dengan menggunakan evaluasi *Silhouette Coefficient* untuk melakukan uji evaluasi terhadap metode *Hierarchical K-means*. Dengan menggunakan *Hierarchical* sebagai penentuan *Centroid* awal dari *K-means*, membuat iterasi *K-means* lebih singkat dan menghasilkan hasil yang baik. Hasil penelitian menunjukkan bahwa tanpa menggunakan normalisasi *Min-max* dan menerapkan 2 *Cluster* menghasilkan *Silhouette Coefficient* dengan nilai 0,646753 yang merupakan struktur terbaik dari berbagai percobaan.

Kata Kunci : Pengelompokan, *Hierarchical K-means Clustering*, *Centroid K-means*

ABSTRACT

Clustering is an Unsupervised Learning method used for grouping data that has not been labeled. K-means is a clustering method that has the ability to classify data relatively quickly, but has difficulty determining the centroid initially. Then the function of this Hierarchical method is to help solve the problem of determining the initial centroid from the K-means method. This study aims to apply the Hierarchical and K-means combination to classify the UASBN SD values and find out how good the evaluation results of the Hierarchical K-means are. The SD UASBN value data used have the attribute values for Bahasa Indonesia, Matematika and IPA. By using the Silhouette Coefficient evaluation to perform an evaluation test of the Hierarchical K-means method. By using Hierarchical as the initial centroid determination of K-means, making K-means iteration shorter and yielding good results. The results showed that without using Min-max normalization and applying 2 clusters, it resulted in a Silhouette Coefficient with a value of 0.646753 which is the best structure of the various experiments.

Keywords : *Grouping, Hierarchical K-means Clustering, Centroid K-means*

