ABSTRACT

One indicator for assessing progress of country is Human Development Index (HDI), as HDI reflects the quality of life and welfare of the population. HDI measures three main dimensions: health, education, and decent standard of living. An approach to analyzing the progress of human development in Indonesia can be done by clustering districts/cities based on HDI values to see differences in the quality of life between regions. The study will apply K-Medoids and Clustering Large Application (CLARA), using Euclidean distance, to group districts/cities in Indonesia. K-Medoids uses distance measurement by selecting one of the data (Medoid) as the cluster center, so it is more robust to outliers. CLARA algorithm works by taking a random sample from the dataset, then applying the K-Medoids algorithm on that sample to determine the optimal medoid. Validation of clustering results is done using Silhouette Coefficient. Higher values indicate more optimal clustering results. Based on Silhouette Coefficient, the best number of clusters for K-Medoids and CLARA is 3 clusters, respectively, with Silhouette Coefficient values of 0,2961 and 0,2853. The results of this study show that the best grouping of districts/cities in Indonesia, based on indicators HDI, is 3 clusters using the K-Medoids algorithm.

Keywords: Human Development Index, Cluster Analysis, K-Medoids, CLARA, Euclidean, Silhouette Coefficient