ABSTRACT

Equitable access to senior high school education in East Kalimantan Province faces challenges in the form of unequal distribution of schools and potential overcapacity in certain areas. This study applies the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) method with Euclidean distance. The study aims to analyze the spatial distribution patterns of 187 public senior high schools in East Kalimantan Province based on geographic coordinate data, namely latitude and longitude. The clustering process begins by determining optimal parameters through exploring the minimum sample value and minimum cluster size. The parameter combination $m_{\text{pts}} = 4$ and $m_{\text{cl}} = 9$ was selected based on the highest and most stable Density-Based Clustering Validation (DBCV) value. The clustering results show the formation of five main clusters accompanied by noise. Cluster performance was measured using DBCV resulting in a value of 0,69823 which indicates good cluster quality. This value also reflects that the model is able to adequately capture spatial patterns based on the geographic distribution of schools. The clusters formed depict areas with relatively close concentrations of schools, while objects identified as noise indicate school locations with low spatial proximity to any cluster and are therefore important to review specifically in the context of equal distribution of educational services.

Keywords: HDBSCAN, DBCV, Overcapacity, Education, Euclidean distance