ABSTRACT

Floods are one of the natural disasters that frequently hit Indonesia, particularly East Java Province. Flood disasters are triggered by various factors such as heavy rainfall, increasing population density, and varying geographical conditions in each region. The impacts of flood disasters are not only economically damaging but also result in loss of life and infrastructure damage. Therefore, accurate data analysis is needed to determine the flood vulnerability levels in each district/city so that mitigation efforts can be carried out more effectively. This study aims to classify districts/cities in East Java Province based on their flood vulnerability levels using the K-Means Clustering method. Additionally, this study evaluates the quality of the clustering results by comparing three types of distance calculations: Euclidean Distance, Manhattan Distance, and Chebyshev Distance. This study uses four main variables in the analysis: average rainfall, population density, area size, and the number of flood incidents during the 2022 period. The clustering results produced four clusters with flood vulnerability levels low (26.32%), moderate (50.00%), high (13.16%), and very high (10.53%). The clustering results were evaluated using the Silhouette Coefficient method to measure how well each data point fits into its respective cluster. The test results showed that the Chebyshev distance calculation yielded the highest Silhouette value of 0.5852, indicating that this method effectively separates clusters and provides more optimal results compared to other distance methods. Furthermore, the clustering results were visualized using Google Looker Studio and a web page publication to facilitate data interpretation and provide an interactive visual representation of the distribution of flood vulnerability levels in East Java. This visualization is expected to serve as a consideration for the government or relevant parties in determining priorities for flood disaster management and mitigation in vulnerable areas.

Keywords— Flood Risk, K-Means, Clustering, Euclidean, Manhattan, Chebyshev