

ABSTRAK

PENGELOMPOKAN DATA RAWAN BENCANA ALAM DI BERBAGAI KOTA DAN KABUPATEN DI INDONESIA BERBASIS ALGORITMA K-MEANS *CLUSTERING*

Oleh

SHERLY MARTINA MULYADI

Indonesia terletak di wilayah cincin api pasifik dan beriklim tropis menyebabkan rawan terhadap bencana alam seperti banjir, gempa bumi, angin puting beliung, letusan gunung api, tanah longsor, dan kekeringan. Bencana alam memberikan dampak kerugian bagi manusia dan lingkungan. Kerugian dapat dikurangi dengan melakukan upaya mitigasi. Salah satu strategi mitigasi, yaitu dengan informasi terkait daerah rawan bencana. Berdasarkan situasi tersebut, penelitian ini memiliki tujuan untuk *clustering* data rawan bencana alam di Indonesia berdasarkan jumlah terjadinya bencana alam sehingga dapat membantu pemangku kepentingan dalam mengidentifikasi daerah rawan bencana. *Clustering* dilakukan menggunakan algoritma *data mining K-Means* dan metode pengembangan *Cross Industry Standard Process for Data Mining (CRISP-DM)*. Jenis bencana alam yang digunakan dalam penelitian ini adalah banjir, tanah longsor, kebakaran hutan dan lahan, gelombang pasang atau abrasi, kekeringan, serta angin puting beliung. Metode *elbow* dimanfaatkan untuk menentukan jumlah *cluster* terbaik. Hasil dari penelitian memperlihatkan bahwa *elbow method* dapat digunakan untuk menentukan jumlah *cluster* yang optimal. Hal ini dibuktikan dengan kekohesifan antara objek dengan *centroid* terbaik terdapat di *cluster* yang dihasilkan oleh model menggunakan atribut kekeringan dengan *silhouette coefficient* sebesar 0,9 dan *davies bouldin index* sebesar 0,39. Penelitian ini menghasilkan lima *cluster*, yaitu *cluster* 0 terdiri atas 391 daerah, *cluster* 1 sebanyak 2 daerah, *cluster* 2 sebanyak 64 daerah, *cluster* 3 sebanyak 25, dan *cluster* 4 sebanyak 11 daerah. Penelitian ini juga membangun visualisasi hasil *clustering* berupa *mapping* daerah rawan bencana alam di Indonesia menggunakan GeoPandas yang mencakup 494 daerah.

Kata kunci : bencana alam, *clustering*, CRISP-DM, *elbow method*, K-Means

ABSTRAK

CLUSTERING OF NATURAL DISASTER PRONE DATA IN VARIOUS CITIES AND DISTRICTS IN INDONESIA BASED ON K-MEANS CLUSTERING ALGORITHM

Oleh

SHERLY MARTINA MULYADI

Indonesia is located in the Pacific Ring of Fire and has a tropical climate, making it prone to natural disasters such as floods, earthquakes, tornadoes, volcanic eruptions, landslides, and droughts. Natural disasters cause losses to humans and the environment. Losses can be reduced by making mitigation efforts. One of the mitigation strategies is information related to disaster-prone areas. Based on this situation, this study aims to cluster data on natural disaster prone areas in Indonesia based on the number of occurrences of natural disasters so that it can help stakeholders in identifying disaster-prone areas. Clustering is done using the K-Means data mining algorithm and the Cross Industry Standard Process for Data Mining (CRISP-DM) development method. The types of natural disasters used in this research are flood, landslide, forest and land fire, tidal wave or abrasion, drought, and tornado. The elbow method was utilized to determine the best number of clusters. The results of the study show that the elbow method can be used to determine the optimal number of clusters. This is evidenced by the cohesiveness between objects with the best centroid in the cluster produced by the model using the drought attribute with a silhouette coefficient of 0.9 and a davies bouldin index of 0.39. This study produced five clusters, namely cluster 0 consisting of 391 regions, cluster 1 as many as 2 regions, cluster 2 as many as 64 regions, cluster 3 as many as 25, and cluster 4 as many as 11 regions. This research also built a visualization of clustering results in the form of mapping natural disaster prone areas in Indonesia using GeoPandas which covers 494 regions.

Keywords: clustering, CRISP-DM, elbow method, K-Means, natural disaster