

INTISARI

Pengelompokan Data Gempa Bumi di Indonesia Menggunakan Metode Reduksi Dimensi t-SNE dan Metode Klastering DBSCAN

Oleh

Atika Indah Mentari

20/459358/PA/20019

Indonesia merupakan negara yang terletak di pertemuan tiga titik lempeng besar bumi dan wilayah Cincin Api Pasifik yang menyebabkan banyak terjadi gempa bumi di Indonesia yang dapat menimbulkan banyak kerugian. Salah satu langkah dalam meminimalisir dampak yang mungkin terjadi adalah dengan mengenali karakteristik gempa bumi di Indonesia melalui pengelompokan data atau klastering. Pada penelitian ini, dilakukan klastering gempa bumi di Indonesia menggunakan metode *density-based spatial clustering of applications with noise* (DBSCAN) pada data gempa bumi di Indonesia tahun 2019 hingga tahun 2023 yang akan dianalisis per tahun. Variabel digunakan adalah *longitude*, *latitude*, *magnitude*, dan *depth*. Pada analisis klastering menggunakan DBSCAN, terdapat metode penentuan parameter ϵ pada DBSCAN dengan melihat *knee point* pada *k-distance plot*. Namun, metode tersebut tidak dapat memberikan hasil yang optimal sehingga diperlukan pemahaman terhadap data melalui penyederhanaan data menggunakan metode reduksi dimensi nonlinear *t-distributed stochastic neighbor embedding* (t-SNE) yang mempermudah visualisasi data, mampu mempertahankan struktur data, dan diharapkan mempermudah penyetelan parameter pada DBSCAN. Hasil dari DBSCAN tanpa menggunakan reduksi dimensi dan menggunakan reduksi dimensi t-SNE dievaluasi menggunakan *average silhouette score* dan indeks Davies Bouldin. Hasil evaluasi menunjukkan bahwa klastering gempa bumi di Indonesia menggunakan metode reduksi dimensi t-SNE pada metode klastering DBSCAN memberikan hasil yang lebih baik daripada tanpa metode reduksi dimensi t-SNE.

Kata kunci: gempa bumi, klastering, DBSCAN, reduksi dimensi, t-SNE

ABSTRACT

Indonesian Earthquake Data Clustering Using Dimensionality Reduction Method t-SNE and Clustering Method DBSCAN

By

Atika Indah Mentari

20/459358/PA/20019

Indonesia is a country located at the intersection of three major tectonic plates and within the Pacific Ring of Fire, which results in frequent earthquakes that can cause significant damage. One of the steps to minimize the potential impact is to understand the characteristics of earthquakes in Indonesia through data clustering. This study aims to cluster earthquakes in Indonesia using the density-based spatial clustering of applications with noise (DBSCAN) method on earthquake data from 2019 to 2023, analyzed annually. The variables to be used are longitude, latitude, magnitude, and depth. In the clustering analysis using DBSCAN, there is a method for determining the parameter ε by examining the knee point on the k-distance plot. However, this method may not provide optimal results, so a better understanding of the data is needed by simplifying the data using the nonlinear dimensionality reduction method t-distributed stochastic neighbor embedding (t-SNE), which facilitates data visualization, preserves data structure, and is expected to assist in tuning DBSCAN parameters. The results of DBSCAN without dimensionality reduction and with t-SNE dimensionality reduction are evaluated using the average silhouette score and Davies Bouldin index. The evaluation results indicate that earthquake clustering in Indonesia using the t-SNE dimensionality reduction method applied to DBSCAN clustering produced better results compared to DBSCAN without t-SNE dimensionality reduction.

Keywords: earthquake, clustering, DBSCAN, dimensionality reduction, t-SNE