

ABSTRAK

Algoritma K-Medoids merupakan salah satu metode klustering yang banyak digunakan dalam data mining untuk mengelompokkan data berdasarkan kemiripan. Namun, pada dataset berukuran besar, algoritma ini sering menghadapi kendala dalam hal prediktabilitas dan stabilitas jumlah iterasi. Penelitian ini bertujuan untuk mengoptimalkan kinerja algoritma K-Medoids dengan menerapkan metode *Z-Score* sebagai pendekatan inisialisasi medoids. Penelitian diawali dengan tinjauan literatur dan perumusan masalah, dilanjutkan dengan pengumpulan data, inisialisasi medoids menggunakan *Z-Score*, penerapan algoritma K-Medoids, serta evaluasi performa menggunakan *Davies-Bouldin Index* (DBI). Pengujian dilakukan terhadap *Whoscale Customer Dataset* dan *QSAR Fish Toxicity Dataset* yang diperoleh dari *UCI Machine Learning Repository*. Hasil penelitian menunjukkan bahwa inisialisasi medoids berbasis *Z-Score* secara signifikan mampu mengurangi jumlah iterasi. Pada *Whoscale Customer Dataset*, jumlah iterasi menurun dari rata-rata 7 menjadi 2, sedangkan pada *QSAR Fish Toxicity Dataset* menurun dari 6 menjadi 2. Evaluasi performa menggunakan DBI juga menunjukkan peningkatan kualitas klaster yang dihasilkan. Nilai rata-rata DBI pada *Whoscale Customer Dataset* menurun dari 2,3210 menjadi 1,2828, sementara pada *QSAR Fish Toxicity Dataset* menurun dari 1,7328 menjadi 0,8768. Penurunan nilai DBI ini mencerminkan terbentuknya klaster yang lebih kompak dan terpisah dengan baik. Temuan ini membuktikan bahwa metode *Z-Score* efektif dalam meningkatkan performa algoritma K-Medoids, khususnya dalam hal efisiensi iterasi dan kualitas klaster.

Kata kunci : *K-Medoids, Z-Score, Klustering, Davies-Bouldin Index (DBI), Optimasi, UCI Machine Learning Repository*

ABSTRACT

The K-Medoids algorithm is one of the clustering methods widely used in data mining to group data based on similarity. However, when applied to large datasets, this algorithm often encounters challenges related to predictability and the stability of iteration counts. This study aims to optimize the performance of the K-Medoids algorithm by implementing the Z-Score method as an approach for medoid initialization. The research process began with a literature review and problem formulation, followed by data collection, medoid initialization using Z-Score, application of the K-Medoids algorithm, and performance evaluation using the Davies-Bouldin Index (DBI). Experiments were conducted on the Whoscale Customer Dataset and the QSAR Fish Toxicity Dataset obtained from the UCI Machine Learning Repository. The findings indicate that Z-Score-based medoid initialization significantly reduces the number of iterations. In the Whoscale Customer Dataset, the average number of iterations decreased from 7 to 2, while in the QSAR Fish Toxicity Dataset, it decreased from 6 to 2. Performance evaluation using DBI also showed an improvement in clustering quality. The average DBI value for the Whoscale Customer Dataset decreased from 2.3210 to 1.2828, and for the QSAR Fish Toxicity Dataset, it decreased from 1.7328 to 0.8768. These lower DBI values indicate the formation of clusters that are more internally cohesive and well-separated. These results demonstrate that the Z-Score method is effective in enhancing the performance of the K-Medoids algorithm, particularly in terms of iteration efficiency and clustering quality.

Keywords: K-Medoids, Z-Score, Clustering, Davies-Bouldin Index (DBI), Optimization, UCI Machine Learning Repositor