

INTISARI

Penelitian ini bertujuan untuk meningkatkan kinerja algoritma K-Means Clustering dengan mengusulkan metode inisialisasi centroid berbasis algoritma hybrid Particle Swarm Optimization dan Grey Wolf Optimizer (PSO-GWO). Metode diimplementasikan dan dievaluasi menggunakan lima dataset berbeda dari UCI Machine Learning Repository. Evaluasi dilakukan dengan mengukur kualitas clustering menggunakan dua metrik utama: Silhouette Score dan Davies-Bouldin Index (DBI). Berdasarkan hasil eksperimen, metode inisialisasi hybrid PSO-GWO menunjukkan peningkatan Silhouette Score yang signifikan dibanding inisialisasi acak dan K-Means++ pada lima dataset, khususnya pada dataset besar dan berdimensi tinggi seperti Covertype dan Secom. Selain itu, PSO-GWO juga terbukti unggul dalam kecepatan konvergensi nilai SSE, DB Index, dan Silhouette Score, sehingga mempercepat proses pencapaian hasil optimal. Hasil uji t berpasangan menunjukkan bahwa perbedaan kinerja PSO-GWO terhadap baseline (K-Means dan K-Means++) signifikan secara statistik ($p < 0.05$), menunjukkan bahwa pendekatan ini dapat menghasilkan struktur cluster yang lebih kompak dan terpisah. Secara keseluruhan, pendekatan PSO-GWO terbukti sebagai alternatif yang menjanjikan untuk mengatasi sensitivitas inisialisasi centroid pada K-Means, terutama pada data yang kompleks.

Kata kunci: K-Means Clustering, Inisialisasi Centroid, PSO-GWO, Metaheuristik, Algoritma Optimasi

ABSTRACT

This research aims to improve the performance of the K-Means clustering algorithm by proposing a centroid initialization method based on a hybrid of Particle Swarm Optimization and Grey Wolf Optimizer (PSO-GWO). The method was implemented and evaluated using five different datasets from the UCI Machine Learning Repository. The evaluation was conducted by measuring clustering quality using two main metrics: Silhouette Score and Davies-Bouldin Index (DBI). Experimental results show that the hybrid PSO-GWO initialization method significantly improves the Silhouette Score compared to random initialization and K-Means++ across all five datasets, particularly on large and high-dimensional datasets such as Covertype and Secom. Furthermore, PSO-GWO also demonstrated superiority in the convergence speed of SSE, DB Index, and Silhouette Score, thereby accelerating the process of reaching optimal results. Paired t-test results indicate that the performance differences between PSO-GWO and the baselines (K-Means and K-Means++) are statistically significant ($p < 0.05$), confirming that this approach produces more compact and well-separated cluster structures. Overall, PSO-GWO proves to be a promising alternative to address the sensitivity of centroid initialization in K-Means, especially for complex datasets.

Keywords: K-Means Clustering, Centroid Initialization, PSO-GWO Hybrid, Metaheuristic, Optimization Algorithm