

ABSTRAK

Usaha Mikro, Kecil, dan Menengah (UMKM) memiliki peran penting dalam mendukung pertumbuhan ekonomi daerah. Namun, proses segmentasi UMKM masih sering dilakukan secara konvensional dan belum optimal dalam menangani *mixed data* yang terdiri atas atribut numerik dan kategorikal. Selain itu, seluruh atribut pada proses *clustering* umumnya diperlakukan memiliki tingkat kepentingan yang sama, padahal setiap atribut memiliki kontribusi informasi yang berbeda terhadap pembentukan *cluster*. Penelitian ini bertujuan untuk meningkatkan kualitas segmentasi UMKM menggunakan metode *Clustering Large Applications (CLARA)* dengan pembobotan dan seleksi fitur berbasis *Entropy weight method (EWM)*. *Dataset* yang digunakan terdiri dari 1.049 data UMKM di Kecamatan Sampang. Metode *CLARA* digunakan dengan perhitungan *Gower distance* untuk menangani *mixed data*, sedangkan *EWM* digunakan untuk pembobotan dan seleksi fitur. Evaluasi *clustering* dilakukan menggunakan *Silhouette Coefficient (SC)* dan *Davies–Bouldin Index (DBI)*. Hasil penelitian menunjukkan bahwa model *baseline* menghasilkan nilai *SC* sebesar 0.4183 dan *DBI* sebesar 1.2393 pada $k = 4$. Setelah diterapkan pembobotan *EWM*, nilai *SC* meningkat sebesar 10.86% menjadi 0.4637 dan *DBI* menurun sebesar 25.29% menjadi 0.9258. Hasil terbaik diperoleh pada skenario *CLARA Top-4* dengan nilai *SC* sebesar 0.5588 dan *DBI* sebesar 0.6181. Dibandingkan *baseline*, nilai *SC* meningkat sebesar 33.59% dan *DBI* menurun sebesar 50.13%. Hasil tersebut menunjukkan bahwa pembobotan dan seleksi fitur berbasis *entropy* mampu meningkatkan kualitas *clustering* UMKM secara lebih optimal dan representatif.

Kata kunci: UMKM, *CLARA*, *Entropy weight*, *Gower Distance*, *Clustering*.

ABSTRACT

Micro, Small, and Medium Enterprises (MSMEs) play an important role in supporting regional economic growth. However, the MSME segmentation process is still often conducted conventionally and is not optimal in handling mixed data consisting of numerical and categorical attributes. In addition, all attributes in the clustering process are generally treated as having the same level of importance, even though each attribute contributes different information to cluster formation. This study aims to improve the quality of MSME segmentation using the Clustering Large Applications (CLARA) method with weighting and feature selection based on the Entropy Weight Method (EWM). The dataset used consists of 1,049 MSME data records from Sampang District. The CLARA method was applied using Gower distance to handle mixed data, while EWM was used for feature weighting and feature selection. Clustering evaluation was conducted using the Silhouette Coefficient (SC) and Davies–Bouldin Index (DBI). The results showed that the baseline model produced an SC value of 0.4183 and a DBI value of 1.2393 at $k = 4$. After applying EWM weighting, the SC value increased by 10.86% to 0.4637, while the DBI value decreased by 25.29% to 0.9258. The best performance was achieved by the CLARA Top-4 scenario with an SC value of 0.5588 and a DBI value of 0.6181. Compared to the baseline model, the SC value increased by 33.59% and the DBI value decreased by 50.13%. These results indicate that entropy-based weighting and feature selection are able to improve the quality of MSME clustering in a more optimal and representative manner.

Keywords: MSMEs, CLARA, Entropy weight, Gower distance, Clustering