

ABSTRAK

Usaha Mikro, Kecil, dan Menengah (UMKM) merupakan sektor penting dalam perekonomian Indonesia, termasuk di Kabupaten Sidoarjo yang memiliki jumlah dan keragaman UMKM yang tinggi. Permasalahan utama dalam pemetaan UMKM adalah ketidakstabilan hasil klasterisasi ketika menggunakan metode *clustering* tunggal. Penelitian ini bertujuan menerapkan *Ensemble Clustering* berbasis *Weighted Voting* untuk menghasilkan pemetaan UMKM yang lebih stabil dan representatif. Data yang digunakan berjumlah 3.320 UMKM dari Dinas Koperasi dan Usaha Mikro Kabupaten Sidoarjo tahun 2025. Penelitian mengikuti kerangka kerja CRISP-DM dengan tahapan data *preparation*, *modeling*, *evaluation*, dan *deployment*. Metode yang digunakan mengombinasikan algoritma K-Means, K-Medoids, dan *Agglomerative Hierarchical Clustering* (AHC), sedangkan pembobotan ditentukan menggunakan fungsi Softmax berdasarkan nilai *Silhouette Score*. Hasil penelitian menunjukkan bahwa nilai cluster optimal adalah K=2 dengan *Silhouette Score* tertinggi sebesar 0,6517, lebih baik dibandingkan K-Means (0,6353), AHC (0,6276), dan K-Medoids (0,5024). Hasil pemetaan mengidentifikasi dua kelompok UMKM, yaitu UMKM Tradisional (96,1%) dan UMKM Digital (3,9%). Temuan ini menunjukkan bahwa *Ensemble Clustering* berbasis *Weighted Voting* mampu menghasilkan pemetaan UMKM yang lebih stabil, konsisten, dan berkualitas dibandingkan metode *clustering* tunggal.

Kata kunci: UMKM, *Ensemble Clustering*, *Weighted Voting*, K-Means, K-Medoids, AHC, *Silhouette Score*.

ABSTRACT

Micro, Small, and Medium Enterprises (MSMEs) are a vital sector of the Indonesian economy, including in Sidoarjo Regency, which has a large number and diverse range of MSMEs. The main challenge in MSME mapping is the instability of clustering results when using a single clustering method. This study aims to implement an Ensemble Clustering approach based on Weighted Voting to produce a more stable and representative MSME segmentation. The dataset consists of 3,320 MSMEs obtained from the Department of Cooperatives and Micro Enterprises of Sidoarjo Regency in 2025. This research follows the CRISP-DM framework, including the stages of data preparation, modeling, evaluation, and deployment. The proposed method combines the K-Means, K-Medoids, and Agglomerative Hierarchical Clustering (AHC) algorithms, while the voting weights are determined using the Softmax function based on Silhouette Score values. The results indicate that the optimal number of clusters is $K = 2$, achieving the highest Silhouette Score of 0.6517, outperforming K-Means (0.6353), AHC (0.6276), and K-Medoids (0.5024). The segmentation identified two MSME groups: Traditional MSMEs (96.1%) and Digital MSMEs (3.9%). These findings demonstrate that the Weighted Voting-based Ensemble Clustering approach produces more stable, consistent, and higher-quality MSME segmentation compared to single clustering methods.

Keywords: MSMEs, Ensemble Clustering, Weighted Voting, K-Means, K-Medoids, Agglomerative Hierarchical Clustering (AHC), Silhouette Score.