

ABSTRAKSI

Usaha Mikro, Kecil, dan Menengah (UMKM) berperan penting dalam perekonomian Indonesia. Di Kabupaten Sampang, UMKM menunjukkan potensi pertumbuhan yang signifikan, dengan besarnya jumlah UMKM terdapat beberapa kendala seperti kurangnya modal usaha dan keterbatasan kemampuan sumber daya manusia. Penelitian ini bertujuan mengelompokkan UMKM untuk membantu Dinas Koperasi, Perindustrian, dan Perdagangan Sampang dalam memberikan rekomendasi kebutuhan yang tepat kepada setiap kelompok untuk pengembangan ekonomi lokal. Metode yang digunakan yaitu Algoritma *K-Means Clustering* yang mengelompokkan data menjadi beberapa *cluster* dengan dibantu metode *elbow* untuk menentukan jumlah *cluster*. Metodologi yang digunakan yaitu CRISP-DM dimulai dari pemahaman bisnis hingga penerapan hasil. Data diolah dengan normalisasi min-max dan label *encoding*. Metode *K-Means* jarak dihitung menggunakan *euclidean distance* kemudian dikelompokkan berdasarkan atribut yaitu jenis usaha, jumlah tenaga kerja, kapasitas produksi, omset, aset, dan surat izin. Hasil uji coba *elbow* menunjukkan titik "siku" berada pada *cluster* 2, selanjutnya evaluasi menggunakan *Davies-Bouldin Index* (DBI) dari rentang $K=2$ hingga $K=10$ menghasilkan cluster terbaik pada $K=2$ dengan nilai DBI terendah 0.4388. dan *silhouette coefficient* dengan nilai tertinggi 0.7388 berada pada $K=2$. Hasil *clustering* menghasilkan dua *cluster*: C1 “Tinggi” dengan jumlah *cluster* 1105 UMKM dan C2 “rendah” dengan jumlah *cluster* 105 UMKM.

Kata Kunci : UMKM, *K-Means Clustering*, CRISP-DM, *elbow*

ABSTRACTION

Micro, Small, and Medium Enterprises (MSMEs) play an important role in the Indonesian economy. In Sampang Regency, MSMEs show significant growth potential, with the large number of MSMEs there are several obstacles such as lack of business capital and limited human resource capabilities. This research aims to categorize MSMEs to assist the Sampang Cooperative, Industry and Trade Office in providing recommendations for the right needs of each group for local economic development. The method used is the K-Means Clustering Algorithm which groups data into several clusters with the help of the elbow method to determine the number of clusters. The methodology used is CRISP-DM starting from understanding the business to applying the results. Data is processed with min-max normalization and label encoding. The K-Means method is calculated using euclidean distance and then grouped based on attributes, namely type of business, number of workers, production capacity, turnover, assets, and licenses. The results of the elbow test show that the “elbow” point is in cluster 2, then evaluation using the Davies-Bouldin Index (DBI) from the range of $K = 2$ to $K = 10$ produces the best cluster at $K = 2$ with the lowest DBI value of 0.4388. and the silhouette coefficient with the highest value of 0.7388 is at $K = 2$. The clustering results produced two clusters: C1 “High” with a cluster number of 1105 MSMEs and C2 “low” with a cluster number of 105 MSMEs.

Keywords: MSMEs, K-Means Clustering, CRISP-DM, elbow