

PERBANDINGAN ALGORITMA *K-MEANS* DAN *AGGLOMERATIVE CLUSTERING* DALAM PENGELOMPOKAN UMKM DI PROVINSI DAERAH ISTIMEWA YOGYAKARTA

Satman Afandi¹, Aris Wahyu Murdiyanto², Ulfi Saidata Aesy³

INTISARI

Latar Belakang: Provinsi Daerah Istimewa Yogyakarta (DIY) memiliki potensi ekonomi kreatif yang besar melalui keberadaan UMKM yang tersebar di berbagai wilayah. Namun, data UMKM yang belum tersaji secara terstruktur dan terintegrasi menyulitkan pemangku kebijakan dan juga wisatawan dalam mengakses informasi yang dibutuhkan. Oleh karena itu, diperlukan pendekatan berbasis data untuk mengelompokkan kategori UMKM secara lebih efektif.

Tujuan: Penelitian ini bertujuan untuk membandingkan efektivitas dua algoritma, yaitu *K-Means* dan *Agglomerative clustering*, dalam mengelompokkan kategori UMKM berdasarkan nama desa dan sektor ekonomi kreatif seperti kuliner, kerajinan tangan, dan seni budaya. Hasil klasterisasi diharapkan dapat memberikan informasi yang lebih terstruktur dan mendukung kebutuhan pemangku kepentingan dan juga wisatawan.

Metode Penelitian: Penelitian ini menggunakan pendekatan CRISP-DM yang terdiri dari enam tahap: pemahaman bisnis, pemahaman data, persiapan data, modeling, evaluasi, dan *deployment*. Proses pengelompokan dilakukan dengan membandingkan algoritma *K-Means* dan *Agglomerative clustering*, serta evaluasi performa menggunakan metrik *Silhouette Score*, *Davies-Bouldin Index*, dan *Calinski-Harabasz Index*.

Hasil: Hasil evaluasi menunjukkan bahwa algoritma *Agglomerative clustering* memberikan hasil klasterisasi yang lebih optimal dibandingkan *K-Means*. *Agglomerative clustering* mampu menggambarkan struktur data kategori UMKM dengan lebih akurat dan stabil, terutama dalam konteks persebaran antar desa. Visualisasi hasil klaster ini juga dinilai mampu membantu wisatawan dalam melihat kategori UMKM yang sesuai dengan preferensi mereka.

Kesimpulan: *Agglomerative clustering* terbukti lebih efektif dalam pengelompokan kategori UMKM berbasis lokasi dan sektor, serta lebih tepat digunakan sebagai dasar visualisasi dalam melihat kategori UMKM berbasis wisata dan ekonomi kreatif di DIY.

Kata-kunci: UMKM, Clustering, *K-Means*, *Agglomerative clustering*, CRISP-DM

¹ Mahasiswa Program Studi S-1 Sistem Informasi Universitas Jenderal Achmad Yani Yogyakarta

² Dosen Program Studi S-1 Sistem Informasi Universitas Jenderal Achmad Yani Yogyakarta

³ Dosen Program Studi S-1 Sistem Informasi Universitas Jenderal Achmad Yani Yogyakarta

COMPARISON OF K-MEANS AND AGGLOMERATIVE CLUSTERING ALGORITHMS IN GROUPING MSMEs IN THE SPECIAL REGION OF YOGYAKARTA

Satman Afandi¹, Aris Wahyu Murdiyanto², Ulfi Saidata Aesy³

ABSTRACT

Background: The Special Region of Yogyakarta (DIY) holds significant creative economy potential through the presence of MSMEs spread across various regions. However, the lack of structured and integrated MSME data makes it difficult for policymakers and tourists to access the information they need. Therefore, a data-driven approach is required to cluster MSMEs more effectively.

Objective: This study aims to compare the effectiveness of two algorithms, namely K-Means and Agglomerative clustering, in grouping MSMEs based on village names and creative economy sectors such as culinary, handicrafts, and cultural arts. The clustering results are expected to provide structured insights that support both tourists and policy makers..

Method: The study applies the CRISP-DM approach, which consists of six stages: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The clustering process was implemented using the Python programming language, and evaluated using Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index.

Result: The evaluation results show that the Agglomerative Clustering algorithm provides a more optimal clustering outcome compared to K-Means. Agglomerative Clustering is able to represent the spatial and sectoral structure of MSMEs more accurately and consistently, particularly in the context of distribution across villages. The visualization of these clustering results is also considered helpful for tourists in identifying MSME categories that match their preferences.

Conclusion: Agglomerative clustering has proven to be more effective in grouping MSMEs based on location and sector, and is more suitable as a foundation for visualizing categories of MSMEs related to tourism and the creative economy in the Special Region of Yogyakarta.

Keywords: MSMEs, Clustering, K-Means, Agglomerative clustering , CRISP-DM

¹ Student of the Information Systems at Jenderal Achmad Yani University Yogyakarta

² Lecturer of the Information Systems at Jenderal Achmad Yani University Yogyakarta

³ Lecturer of the Information Systems at Jenderal Achmad Yani University Yogyakarta