

ABSTRAK

Pengelompokkan desa berdasarkan produktivitas pertanian krusial untuk pengelolaan sektor pertanian. Studi ini membandingkan metode clustering, K-Means dan K-Medoids, menggunakan data sembilan variabel termasuk luas tanaman, varietas, umur tanaman, organisme, pupuk, luas terserang, intensitas, luas waspada, dan hasil panen. Tantangan utama adalah jumlah desa yang besar, mempersulit pengelompokan manual oleh Dinas Pertanian. K-Means dan K-Medoids diaplikasikan untuk mengatasi kompleksitas tersebut. Data diolah melalui Label encoding, Imputasi Data, dan Normalisasi Data dengan Min-Max. Eksperimen dilakukan dengan evaluasi SSE, dan hasilnya menunjukkan K-15 terbaik untuk K-Means dan K-14 terbaik untuk K-Medoids, dengan nilai SSE terendah masing-masing 24.0026 dan 35.3600. Rekomendasi Dinas Pertanian Sumenep adalah menggunakan 3 cluster dengan detail C1 produktivitas hasil panen tinggi, C2 produktivitas hasil panen sedang, dan C3 produktivitas hasil panen rendah. K-Means dipilih karena SSE lebih rendah, 62.5151, dibandingkan K-Medoids, 82.2660. Analisis cluster pada K-3 menunjukkan K-Means lebih optimal, dengan presentase hasil panen sesuai kategorinya. Keputusan akhir mengakomodasi permintaan pihak berwenang, memilih K-Means dengan 3 kluster sebagai solusi untuk analisis data hasil panen Sumenep. Penelitian ini berkontribusi dalam pengelompokan desa berdasarkan produktivitas pertanian, mendukung sektor pertanian dalam pengambilan keputusan yang efektif.

Kata kunci : *Clustering, K-Means, K-Medoids, Sum of Squared Errors (SSE)*

ABSTRACT

The clustering of villages based on agricultural productivity is crucial for managing the agricultural sector. This study compares two clustering methods, K-Means and K-Medoids, using data from nine variables, including crop area, varieties, crop age, organisms, fertilizers, affected area, intensity, alert area, and harvest yield. The main challenge lies in the large number of villages, making manual grouping by the Agriculture Department difficult. K-Means and K-Medoids are applied to address this complexity. The data is processed through Label Encoding, Data Imputation, and Data Normalization using Min-Max. Experiments are conducted with SSE evaluation, and the results indicate that K-15 is optimal for K-Means and K-14 is optimal for K-Medoids, with respective lowest SSE values of 24.0026 and 35.3600. The Agriculture Department of Sumenep recommends using 3 clusters with details of C1 high crop yield productivity, C2 medium crop yield productivity, and C3 low crop yield productivity. K-Means is chosen because it has a lower SSE of 62.5151 compared to K-Medoids, which has an SSE of 82.2660. Cluster analysis on K-3 shows that K-Means is more optimal, with harvest yield percentages matching their categories. The final decision accommodates the authorities' request, selecting K-Means with 3 clusters as the solution for Sumenep's harvest data analysis. This research contributes to village grouping based on agricultural productivity, supporting the agricultural sector in effective decision-making.

Keywords: *Clustering, K-Means, K-Medoids, Sum of Squared Errors (SSE).*