

ABSTRAK

Penelitian ini bertujuan untuk membandingkan dua algoritma clustering, yaitu *K-Means* dan *K-Medoids*, dalam mengelompokkan siswa pada Sekolah Menengah Kejuruan Leonardo Klaten yang memiliki minat untuk berkuliah di program studi informatika. Data yang digunakan dalam penelitian ini adalah hasil ujian akademik siswa kelas XII tahun 2023 yang terdiri dari 249 siswa dengan berbagai mata pelajaran yang diuji. Proses penelitian dimulai dengan tahap pengumpulan data, *preprocessing* untuk membersihkan dan menstandarisasi data, kemudian dilanjutkan dengan penerapan kedua algoritma clustering tersebut. Evaluasi dilakukan dengan menggunakan metode *Elbow* untuk menentukan jumlah *cluster* optimal dan *Silhouette Score* untuk mengukur kualitas cluster yang terbentuk. Hasil penelitian menunjukkan bahwa algoritma *K-Means* memberikan hasil yang lebih baik dibandingkan dengan *K-Medoids*, dengan nilai *Silhouette Score* yang lebih tinggi, serta pemisahan *cluster* yang lebih jelas pada visualisasi PCA. *K-Means* berhasil mengelompokkan 232 siswa yang menunjukkan minat untuk berkuliah di Informatika, sementara 17 siswa lainnya tidak memiliki minat tersebut. Penelitian ini memberikan kontribusi pada literatur mengenai teknik *clustering* dalam bidang pendidikan, khususnya dalam mengidentifikasi siswa yang berpotensi melanjutkan pendidikan tinggi di program studi informatika.

Kata kunci: *K – Means*, *K – Medoids*, Clustering, Sekolah Menengah Kejuruan



ABSTRACT

This research aims to compare two clustering algorithms, namely K-Means and K-Medoids, in clustering students at Vocational High School Leonardo in Klaten who have an interest in studying informatics. The data used in this study is the academic exam results of 12th-grade students in 2023, consisting of 249 students across various subjects. The research process started with data collection, followed by preprocessing to clean and standardize the data, and then applying both clustering algorithms. The evaluation was conducted using the Elbow method to determine the optimal number of clusters and Silhouette Score to assess the quality of the clusters formed. The results of the study showed that the K-Means algorithm performed better than K-Medoids, with a higher Silhouette Score and clearer cluster separation in the PCA visualization. K-Means successfully clustered 232 students who expressed interest in pursuing a degree in informatics, while 17 other students did not show such interest. This research contributes to the literature on clustering techniques in the field of education, particularly in identifying students with potential to pursue higher education in informatics.

Keywords: K – Means, K – Medoids, Clustering, Informatics, Vocational High School

