

Abstrak

Penelitian ini bertujuan untuk mengimplementasikan algoritma K-Means *Clustering* dalam melakukan segmentasi pasien berdasarkan riwayat kunjungan dan diagnosis penyakit. *Dataset* yang digunakan berasal dari Kaggle yang berisi data pasien dengan atribut usia, kondisi medis, jenis pelayanan, biaya tagihan, dan lama rawat inap. Proses penelitian meliputi tahap *preprocessing*, transformasi data, normalisasi, serta penerapan algoritma K-Means menggunakan aplikasi KNIME. Jumlah cluster yang digunakan sebanyak tiga cluster. Hasil penelitian menunjukkan bahwa data pasien berhasil dikelompokkan menjadi tiga kategori, yaitu pasien dengan kondisi ringan, sedang, dan berat. *Cluster* dengan kondisi berat memiliki jumlah anggota terbanyak yang menunjukkan adanya pasien dengan kebutuhan penanganan lebih intensif. Hasil *clustering* ini mampu memberikan gambaran pola segmentasi pasien sehingga dapat digunakan sebagai dasar dalam mendukung pengambilan keputusan serta meningkatkan kualitas pelayanan kesehatan.

Abstract

This study aims to implement the K-Means Clustering algorithm to segment patients based on visit history and medical diagnosis. The dataset used was obtained from Kaggle, containing patient data such as age, medical condition, admission type, billing amount, and length of stay. The research process includes data preprocessing, transformation, normalization, and the implementation of the K-Means algorithm using the KNIME application. The number of clusters used in this study is three. The results show that patient data can be grouped into three categories: low, medium, and high condition levels. The cluster with high condition levels has the largest number of members, indicating patients who require more intensive care. These clustering results provide insights into patient segmentation patterns and can be used to support decision-making and improve healthcare service quality.