

ABSTRACT

Data Mining is also the process of extracting patterns and information from large data sets. One of the Data Mining techniques used in this study is K-Means Clustering, which is a clustering method that divides data into groups based on common characteristics. Clustering works by determining the centroid of each cluster and grouping the data based on the closest distance to that centroid. This study aims to classify the data of family planning patients into short-term family planning and long-term family planning based on the characteristics of the name, age, and number of children. The Data used were 50 KB patient data in 2023 analyzed using RapidMiner software which facilitates faster and more accurate data processing. The K-Means Clustering algorithm groups patients into 2 clusters, namely cluster c0 which states that patients are recommended to use short-term KB types, and cluster C1 which recommends long-term KB types. The results of this study showed that cluster c0 consists of 29 members, while cluster C1 has 21 members. The use of K-Means Clustering algorithm is able to effectively divide patient data that provides benefits to midwives in providing recommendations for more targeted and structured types of family planning to patients. RapidMiner also provides consistent results, and simplifies the data analysis process and ensures that manual clustering results are in line.

Keywords: K-Means Clustering, Data Mining, RapidMiner, Contraception, Puskesmas

ABSTRAK

Data Mining juga menjadi proses untuk mengekstrak pola dan informasi dari kumpulan data yang besar. Salah satu teknik Data Mining yang digunakan pada penelitian ini adalah *K-Means Clustering*, yang merupakan metode pengelompokan yang membagi data dalam kelompok berdasarkan kesamaan karakteristik. *Clustering* bekerja dengan menentukan *centroid* dari setiap *cluster* dan mengelompokkan data berdasarkan jarak terdekat ke *centroid* tersebut. Penelitian ini bertujuan untuk mengelompokkan data pasien KB kedalam jenis KB jangka pendek dan jenis KB jangka panjang berdasarkan karakteristik nama, umur, dan jumlah anak. Data yang digunakan yaitu 50 data pasien KB tahun 2023 dianalisis menggunakan *software* RapidMiner yang memfasilitasi pengolahan data lebih cepat dan akurat. Algoritma *K-Means Clustering* mengelompokkan pasien ke dalam 2 *cluster*, yaitu *cluster* C0 yang menyatakan pasien direkomendasikan menggunakan jenis KB jangka pendek, dan *cluster* C1 yang merekomendasikan jenis KB jangka panjang. Hasil dari penelitian ini menunjukkan bahwa *cluster* C0 terdiri dari 29 anggota, sedangkan *cluster* C1 memiliki 21 anggota. Penggunaan algoritma *K-Means Clustering* mampu secara efektif membagi data pasien yang memberikan manfaat kepada bidan dalam memberikan rekomendasi jenis KB yang lebih tepat sasaran dan terstruktur kepada pasien. RapidMiner juga memberikan hasil yang konsisten, dan mempermudah proses analisis data dan memastikan bahwa hasil *clustering* manual sejalan.

Kata Kunci: K-Means Clustering, Data Mining, RapidMiner, Kontrasepsi, Puskesmas