

Penerapan Algoritma K-Means Clustering untuk Mengelompokkan Rekomendasi Metode Kontrasepsi Berbasis Machine Learning

Quratih Adawiyah[✉], Sarjon Defit, Sumijan

Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK Padang, 25221, Indonesia

quratih.adawiyah@gmail.com

Abstract

Data mining is also a process to extract patterns and information from large data sets. One of the techniques used in this study is K-Means Clustering which is a clustering method that divides data into groups based on similar characteristics. Clustering works by identifying the centroid of each cluster and grouping the data based on the closest distance to the centroid. This study aims to provide recommendations for appropriate contraceptive methods for fertile couples by grouping family planning patient data based on characteristics such as name, age, and number of children. The K-Means Clustering algorithm is used to group recommendations for contraceptive methods based on patient data collected to obtain more accurate recommendations. The data used are name data, age data, and number of children of family planning patients in 2023 at the Puskesmas. The dataset used is 20 data. The methods used include literature studies, data collection through observation, and interviews. The results of this study can group data into 2 clusters, namely cluster C0 states the short-term method while C1 states the long-term method. From the results of the clustering process, the position of the last cluster in iteration 4 does not change anymore, then the process is declared complete, and short-term cluster C0 with 11 members and C1 with 9 members. This study also utilizes the RapidMiner application to assist the clustering process. The results show that this algorithm is effective in identifying patterns and groups in family planning patient data, which is expected to help health workers provide more accurate recommendations and contribute to improving health services, especially in family planning services at the Community Health Center.

Keywords: Artificial Intelligence, Machine Learning, Data Mining, K-Means Clustering, RapidMiner

Abstrak

Data mining juga menjadi proses untuk mengekstrak pola dan informasi dari kumpulan data yang besar. Salah satu teknik yang digunakan pada penelitian ini adalah K-Means Clustering yang merupakan metode pengelompokan yang membagi data dalam kelompok berdasarkan kesamaan karakteristik. Clustering bekerja dengan mengidentifikasi centroid dari setiap cluster dan mengelompokkan data berdasarkan jarak terdekat ke centroid tersebut. Penelitian ini bertujuan untuk memberikan rekomendasi metode kontrasepsi yang tepat kepada pasangan usia subur dengan mengelompokkan data pasien KB berdasarkan karakteristik seperti nama, umur, dan jumlah anak. Algoritma K-Means Clustering digunakan untuk mengelompokkan rekomendasi metode kontrasepsi berdasarkan data pasien yang dikumpulkan sehingga mendapatkan rekomendasi yang lebih akurat. Data yang digunakan adalah data nama, data umur dan jumlah anak pasien KB pada tahun 2023 di Puskesmas. Dataset yang digunakan sebanyak 20 data. Metode yang digunakan meliputi studi literatur, pengumpulan data melalui observasi dan wawancara. Hasil dari penelitian ini dapat mengelompokkan data menjadi 2 cluster, yaitu cluster C0 menyatakan metode jangka pendek sedangkan C1 menyatakan metode jangka panjang. Dari hasil proses cluster, posisi cluster terakhir pada iterasi 4 dan tidak berubah lagi maka dinyatakan proses telah selesai, didapatkan cluster Jangka pendek C0 sebanyak 11 anggota dan C1 berjumlah 9 anggota. Penelitian ini juga memanfaatkan aplikasi RapidMiner untuk membantu proses clustering. Hasil menunjukkan bahwa algoritma ini efektif dalam mengidentifikasi pola dan kelompok dalam data pasien KB, yang diharapkan dapat membantu tenaga kesehatan dalam memberikan rekomendasi yang lebih akurat, dan memberikan kontribusi dalam peningkatan pelayanan kesehatan, khususnya dalam pelayanan KB di Puskesmas.

Keywords: Artificial Intelligence, Machine Learning, Data Mining, K-Means Clustering, RapidMiner

KomtekInfo is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



1. Pendahuluan

Knowledge Discovery in Database (KDD) adalah proses analisa yang terstruktur bertujuan mendapatkan informasi yang baru dan benar, menemukan pola dari data yang kompleks, dan bermanfaat [1]. Pada tahapan KDD terdapat proses Data Mining, yaitu melakukan analisa dari sejumlah besar kumpulan data observasi [2]. Data Mining merupakan metode yang digunakan

untuk mengidentifikasi pola-pola tertentu dari sejumlah besar data, termasuk berbagai teknik seperti teknik klasifikasi [3]. *Clustering* adalah sebuah metode pengelompokan data berdasarkan informasi yang didapatkan dari data yang menguraikan objek tersebut [4]. Pada metode *clustering* terdapat beberapa algoritma salah satunya adalah algoritma *K-Means Clustering* [5]. *K-Means Clustering* merupakan salah satu teknik pengelompokan data dengan cara membagi data yang

ada kedalam bentuk satu atau lebih *cluster* [6]. Karena kesederhanaan memiliki Tingkat ketelitian yang tinggi terhadap ukuran data sehingga K-Means Clustering lebih efisien dan terukur dalam pengolahan data berjumlah besar, selain itu urutan objek tidak mempengaruhi algoritma ini [7].

Menurut undang-undang nomor 52 tahun 2009 tentang perkembangan kependudukan dan Pembangunan keluarga disebutkan bahwa keluarga berencana adalah upaya pengaturan kelahiran anak, jarak kelahiran yang ideal, usia melahirkan yang ideal, dan pengaturan kehamilan melalui promosi, perlindungan dan dukungan sesuai dengan hak reproduksi, guna membangun keluarga yang berkualitas [8].

Berdasarkan penelitian yang dilakukan oleh Ula dkk tentang Penerapan Machine Learning Clustering K-Means dan Linear Regression dalam Penentuan Tingkat Resiko Tuberkulosis Paru, hasil yang diperoleh dalam pengklusteran daerah rawan tuberkulosis paru yang terdapat 2 daerah yang tergolong ke dalam cluster 1. Kemudian 6 daerah tergolong ke dalam cluster 2 dan 9 daerah tergolong ke dalam cluster 3. Hasil prediksi dengan algoritma Regresi Linier 0,5740. Hasil prediksi dari penelitian ini berpengaruh terhadap variabel lain dengan nilai 9,4814. Hasil penelitian ini dapat dijadikan rujukan dinas kesehatan dalam menindaklanjuti penyakit paru [9].

Penelitian juga dilakukan oleh Putra Aryadi dkk tentang Penerapan Algoritma K-Means Untuk Melakukan Klasterisasi Pada Varietas Padi hasil yang diperoleh bahwa lima kelompok cluster varietas padi berhasil diidentifikasi melalui proses klasterisasi. Dari hasil pengujian terdapat nilai Davies Bouldin Index sebesar 0,48, yang menandakan kualitas cluster yang baik [10].

Selanjutnya penelitian yang dilakukan Pane dkk tentang Implementasi Data Mining dengan K-Means Clustering untuk Memprediksi Pengadaan Obat, hasil yang diperoleh bahwa dari 28 data uji yang digunakan didapatkan hasil cluster pertama yang terdiri dari 24 item dengan nilai Davies Bouldin Index yang didapatkan sebesar 0,276 [11].

Selanjutnya penelitian yang dilakukan Sari dkk tentang Penerapan Data Mining Dalam Pengelompokan Buku Yang Dipinjam Menggunakan Algoritma K-Means yang diperoleh bahwa Algoritma *K-Means* digunakan untuk membantu dalam pengambilan keputusan di masa mendatang. Setiap data dipindahkan ke *cluster* berdasarkan kedekatannya dengan *centroid*. *Cluster I* memiliki data dengan jumlah buku yang lebih tinggi, sedangkan *Cluster II* memiliki data dengan jumlah buku yang lebih rendah [12].

Penelitian juga dilakukan Purba dkk tentang Penerapan Data Mining Untuk Pengelolaan Data Rekam Medis Menggunakan Metode *K-Means Clustering* Pada

Rumah Sakit Royal Prima Medan yang diperoleh data rekam medis dapat dikelompokkan menjadi 4 *cluster* dengan 2 variabel, yaitu : jenis kelamin dan jenis perawatan dan hasil analisis ini memperoleh hasil *cluster 1* sebanyak 18217 pasien, *cluster 2* sebanyak 1016 pasien, *cluster 3* sebanyak 396 pasien dan *cluster 4* sebanyak 307 pasien yang di rawat inap pada Rumah Sakit Royal Prima Medan. Peneliti dapat menemukan informasi baru tentang bagaimana RS royal prima medan mengelola berbagai jenis perawatan [13].

Selanjutnya penelitian yang dilakukan oleh Riana dkk tentang identifikasi citra pap smear dengan menggunakan K-Means Clustering, hasil yang diperoleh menunjukkan akurasi yang baik dengan rata-rata akurasi thinprep sebesar 93,33% dan nonthinprep sebesar 90% hasil ini penting untuk deteksi dini kanker serviks [14]. Selanjutnya penelitian yang dilakukan oleh Amijoyo dkk tentang penerapan metode algoritmapengelompokan K-Means pada penerimaan mahasiswa baru, hasil yang diperoleh melibatkan beberapa iterasi dan perhitungan untuk menentukan kelompok akhir. Hasilnya menunjukkan efektivitas algoritma *K-Means Clustering* dalam pengelompokan data siswa untuk tujuan pengambilan keputusan [15].

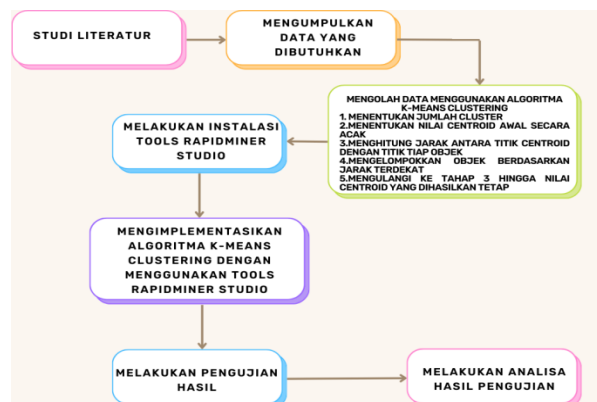
Sebagai salah satu Puskesmas menghadapi tantangan utama untuk mengelompokkan rekomendasi jenis KB yang tepat kepada pasien KB berdasarkan karakteristik dan kebutuhan mereka. Penerapan algoritma K-Means Clustering untuk mengelompokkan rekomendasi metode kontrasepsi berbasis machine learning memiliki keunikan karena fokusnya memberikan rekomendasi yang lebih akurat kepada pasangan usia subur dalam memilih metode kontrasepsi melalui pengelompokan data pasien yang efektif. Selain itu, penelitian ini dapat membantu tenaga kesehatan khususnya bidan, dalam memberikan pelayanan yang lebih baik di Puskesmas. Penerapan algoritma K-Means Clustering, diharapkan dapat meningkatkan efisiensi dan efektivitas pengelolaan data dalam pelayanan keluarga berencana.

Berdasarkan latar belakang yang telah diuraikan, penulis mempunyai tujuan penelitian yaitu untuk menerapkan algoritma K-Means Clustering dalam mengelompokkan rekomendasi metode kontrasepsi berbasis machine learning di Puskesmas. Selain itu, penelitian ini bertujuan untuk membantu tenaga kesehatan, khususnya bidan dalam meningkatkan kualitas pelayanan keluarga berencana melalui pengelompokan data pasien yang efektif dan efisien.

2. Metodologi Penelitian

Pendekatan metodologi dalam penelitian ini difokuskan pada penggunaan algoritma K-Means clustering untuk mengelompokkan data yang telah dikumpulkan. Proses ini dimulai dengan mengumpulkan dan mempersiapkan data, kemudian diikuti dengan penerapan algoritma

menggunakan tools RapidMiner Studio. Tahapan akhir melibatkan pengujian hasil clustering dan analisis untuk menilai efektivitas serta keakuratan metode yang digunakan dalam penelitian ini. Metodologi adalah cara yang digunakan dalam memperoleh berbagai data untuk diproses menjadi informasi yang lebih akurat sesuai permasalahan yang akan diteliti. Kerangka kerja dari penelitian ini adalah untuk mencapai tujuan yakni mengelompokkan rekomendasi metode kontrasepsi. Bentuk tahapan bisa dilihat pada Gambar 1.



Gambar 1 Kerangka Penelitian

Gambar di atas menunjukkan kerangka kerja penelitian yang menjelaskan tahapan-tahapan utama dalam proses penelitian ini. Setiap tahap, mulai dari studi literatur hingga pengujian dan analisis hasil, disusun secara sistematis untuk mencapai tujuan penelitian dengan menggunakan alat bantu RapidMiner Studio. Kerangka ini memberikan panduan langkah demi langkah untuk memastikan bahwa setiap elemen penelitian saling berhubungan dan mendukung keseluruhan proses. Tahapan pada kerangka kerja sebagai berikut:

1. Studi Literatur

Literatur yang digunakan dalam penelitian ini berupa jurnal-jurnal ilmiah Data Mining tentang algoritma K-Means Clustering, serta artikel pendukung lainnya dan dijadikan referensi untuk membantu proses penelitian ini. Studi literatur dalam penelitian ini dilakukan untuk memahami konsep dasar dan teori yang mendasari penggunaan algoritma K-Means clustering. Dengan mengkaji berbagai sumber ilmiah, penelitian sebelumnya, dan referensi terkait, peneliti dapat mengidentifikasi metode dan teknik yang paling relevan untuk diterapkan. Hasil dari studi literatur ini menjadi landasan kuat dalam merancang kerangka kerja penelitian dan menentukan langkah-langkah yang akan diambil.

2. Mengumpulkan Data yang Dibutuhkan

Pada tahap ini tahap mengumpulkan data dilakukan dengan dua cara yaitu observasi dan wawancara. Observasi yaitu pengumpulan data dengan melakukan pengamatan langsung di Puskesmas Pegang Baru, data yang dikumpulkan merupakan data pasien KB pada tahun 2023. Sedangkan wawancara dengan Bidan

pelayanan KB dilakukan untuk mendapatkan informasi tentang keluarga berencana (KB).

Pengumpulan data yang dibutuhkan merupakan langkah awal yang sangat penting dalam penelitian ini. Data dikumpulkan dari berbagai sumber yang relevan dan diseleksi berdasarkan kriteria tertentu untuk memastikan kualitas dan kesesuaiannya dengan tujuan penelitian. Setelah data terkumpul, langkah selanjutnya adalah melakukan proses pre-processing untuk membersihkan dan menyiapkan data sebelum digunakan dalam analisis lebih lanjut.

3. Mengolah Data Menggunakan Algoritma K-Means Clustering

a. Menentukan jumlah cluster

Pada tahap ini peneliti akan menggunakan algoritma K-Means Clustering untuk membagi data menjadi 2 cluster. Cluster C0 menyatakan pasien KB direkomendasikan menggunakan jenis KB jangka pendek, sedangkan C1 menyatakan jenis KB jangka panjang. Menentukan jumlah cluster merupakan tahap krusial dalam penerapan algoritma K-Means clustering, karena hal ini mempengaruhi hasil pengelompokan data. Jumlah cluster yang optimal biasanya ditentukan melalui metode evaluasi seperti Elbow Method atau Silhouette Score, yang membantu mengidentifikasi titik di mana penambahan cluster tidak lagi memberikan peningkatan yang signifikan. Dengan menentukan jumlah cluster yang tepat, peneliti dapat memastikan bahwa data dikelompokkan secara efektif, mencerminkan pola dan hubungan yang ada di dalamnya.

b. Menentukan nilai centroid awal secara acak

Penentuan nilai centroid awal secara acak adalah langkah pertama dalam proses K-Means clustering yang mempengaruhi hasil akhir pengelompokan data. Centroid awal ini digunakan sebagai titik awal untuk mengelompokkan data ke dalam cluster yang berbeda. Setelah centroid awal ditentukan, algoritma akan iteratif memperbaiki posisi centroid hingga konvergensi tercapai, memastikan bahwa data dikelompokkan dengan optimal sesuai dengan pola yang ada. Dalam melakukan penentuan nilai centroid awal ini dapat diambil secara acak dari data sampel yang digunakan dalam penelitian. Pada penelitian ini, nilai centroid untuk C0 diambil dari data sampel ke 2 dan C1 diambil dari data sampel ke 5. Rumus untuk menyatakan kekuatan ini disajikan pada Persamaan 1.

$$\bar{v}_{ij} = \frac{1}{N_i} \sum_{k=10}^{N_i} x_{kj} \quad (1)$$

c. Menghitung jarak antara titik centroid dengan titik tiap objek

Pada tahapan ini dilakukan perhitungan terhadap pusat *cluster* dari titik data dan dihitung jarak masing-masing data ke pusat *cluster* yang terdekat. Dalam menghitung jarak antara data dengan pusat *cluster* tersebut dapat

menggunakan rumus jarak *Euclidean Distance* seperti pada pembahasan sebelumnya. Menggunakan rumus *Euclidean Distance* pada Persamaan 2.

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (2)$$

Mengelompokkan objek berdasarkan jarak terdekat

- d. Mengulangi ke tahap 2 hingga nilai centroid yang dihasilkan tetap dan anggota cluster tidak berpindah ke cluster lain.

Jika pada proses ini pusat *cluster* tidak terjadi perubahan lagi maka proses *clustering* dapat diselesaikan. Jika masih mengalami perubahan maka dapat kembali ke langkah sebelumnya sampai pusat *cluster* tidak terjadi perubahan. Selain hal ini *clustering* juga dapat dihentikan dengan menggunakan batasan maksimal terhadap suatu proses *clustering*.

4. Melakukan Installasi Tools RapidMiner

Tujuan dari installasi tools adalah untuk menguji hasil dari data yang telah diolah dengan menggunakan Microsoft Excel 2021 dengan menggunakan algoritma K-Means Clustering. Dengan tujuan, apakah data yang dihasilkan sebelumnya dengan data yang akan diolah menggunakan tools RapidMiner Studio memiliki hasil yang sama.

5. Mengimplementasikan Algoritma K-Means Clustering dengan Menggunakan Tools RapidMiner

Setelah hasil dari data yang diolah diperoleh dan dianalisa, selanjutnya data tersebut dilakukan dengan menggunakan bantuan tools RapidMiner dan dengan menerapkan langkah-langkah dari algoritma K-Means Clustering. Sehingga dapat menghasilkan hasil yang sesuai dengan yang diharapkan. Dan hasilnya dapat membantu tenaga kesehatan khususnya Bidan pada pelayanan KB untuk mengelompokkan rekomendasi jenis KB.

6. Melakukan Pengujian Hasil

Adapun beberapa tahapan-tahapan yang dilakukan dalam melakukan pengujian adalah:

- a. Melakukan pengujian menggunakan *Microsoft Excel 2021* dengan algoritma *K-Means Clustering*
 - b. Melakukan pengujian menggunakan algoritma *K-Means Clustering* dengan bantuan *software RapidMiner Studio*
 - c. Hasil yang diperoleh memiliki tingkat akurasi yang sesuai dengan yang diharapkan
- #### 7. Melakukan Analisa Hasil Pengujian

Analisa hasil pengujian bertujuan untuk mengevaluasi keakuratan dan efektivitas model clustering yang telah diterapkan. Dalam analisis ini, berbagai metrik seperti *Silhouette Score* atau *Davies-Bouldin Index* digunakan untuk menilai

kualitas pengelompokan data. Hasil analisa ini membantu peneliti memahami sejauh mana model dapat mengidentifikasi pola dalam data serta menentukan apakah model tersebut memenuhi tujuan penelitian. Melakukan analisa hasil yang telah diperoleh dari data yang telah diolah menggunakan algoritma K-Means Clustering berdasarkan permasalahan yang terjadi di Puskesmas.

3. Hasil dan Pembahasan

3.1. Menentukan jumlah cluster

Menentukan jumlah cluster dalam algoritma K-Means adalah langkah penting yang mempengaruhi akurasi dan interpretasi hasil pengelompokan data. Biasanya, metode seperti *Elbow Method* atau *Gap Statistic* digunakan untuk menentukan jumlah cluster yang optimal, dengan tujuan menemukan keseimbangan antara jumlah cluster dan variasi data yang dijelaskan. Pemilihan jumlah cluster yang tepat membantu dalam mengidentifikasi pola atau segmen yang jelas dalam data, sehingga hasil analisis lebih bermakna dan dapat diandalkan. Pada tahap ini peneliti membagi data menjadi 2 cluster yaitu cluster C0 menyatakan metode kontrasepsi jangka pendek dan cluster C1 menyatakan metode kontrasepsi jangka panjang.

3.2. Menentukan Nilai Centroid Secara Acak

Penentuan nilai centroid secara acak merupakan langkah awal dalam algoritma K-Means clustering yang berfungsi sebagai titik awal untuk mengelompokkan data ke dalam cluster yang berbeda. Nilai centroid ini diambil dari sampel data yang ada dan akan digunakan sebagai acuan dalam proses iterasi untuk memperbarui posisi centroid hingga proses konvergensi tercapai. Pemilihan nilai centroid secara acak memungkinkan algoritma untuk menjelajahi berbagai kemungkinan pengelompokan data secara efektif. Penentuan nilai centroid awal dapat diambil secara acak dari 20 data pasien KB yang akan digunakan dalam penelitian. Centroid untuk C0 diambil dari data sampel ke 2 yaitu (32, 3) dan C1 diambil dari data sampel ke 5 yaitu (48, 7).

3.3. Perhitungan *Euclidian Distance*

Perhitungan *Euclidean Distance* digunakan untuk mengukur jarak antara dua titik data dalam ruang multidimensi. Jarak ini dihitung dengan mengkuadratkan selisih setiap pasangan koordinat, menjumlahkan hasilnya, dan mengambil akar kuadrat dari jumlah tersebut. *Euclidean Distance* digunakan dalam K-Means clustering untuk menentukan kedekatan data dengan centroid, sehingga data dapat dikelompokkan ke dalam cluster yang tepat. Pada tahap ini menggunakan rumus 2, adapun hasilnya disajikan pada Tabel 1.

Tabel 1 Hasil iterasi 1

Nama	C0	C1	Cluster
------	----	----	---------

Sri Wahyuni	2,236067977	14,86606875	C0
Endang	0	16,4924225	C0
Irma Khairani	8,246211251	8,246211251	C0
Hotlina	10,19803903	6,32455532	C1
Sahriani	16,4924225	0	C1
Romauli	7	9,848857802	C0
Titi Namara	9,055385138	7,615773106	C1
Yulia	9,055385138	7,615773106	C1
Siti Sumanti	17,2626765	1,414213562	C1
Fitria Delmayani	11,18033989	5,385164807	C1
Yuni Hesti	5,385164807	21,84032967	C0
Siti Maryam	2,828427125	18,97366596	C0
Risna Yanti	8,246211251	24,73863375	C0
Masdaniah	1	16,76305461	C0
Ratna Arianti	4,123105626	13	C0
Febri Yanti	8,062257748	8,544003745	C0
Yusmawarti	3	13,60147051	C0
Arianti	3	13,60147051	C0

Tabel 1 menunjukkan hasil iterasi pertama dari proses K-Means clustering, di mana setiap objek data dikelompokkan ke dalam cluster berdasarkan jarak terdekat dengan centroid. Pada iterasi ini, terdapat dua cluster, yaitu C0 dan C1, dengan objek yang memiliki jarak terdekat ke centroid masing-masing cluster. Tabel 1 menunjukkan hasil iterasi 1. Terdapat cluster C0 sebanyak 12 anggota. Sedangkan cluster C1 sebanyak 8 anggota. Setelah itu didapatkan titik centroid 2, dapat dilihat pada Tabel 2 sebagai berikut.

C0	33,66666667	2,5
C1	44,75	5,375

Tabel titik centroid 2 menunjukkan posisi centroid yang diperbarui setelah iterasi pertama dari proses K-Means clustering. Tabel 2 menunjukkan titik centroid C0 untuk umur yaitu 33,66666667. Sedangkan untuk jumlah anak titik centroid C0 yaitu 2,5. Titik centroid C1 untuk umur yaitu 44,75. Sedangkan untuk jumlah anak yaitu 5,375.

3.4. Mengelompokkan objek untuk menentukan anggota cluster berdasarkan jarak

Pada tahap ini dapun hasilnya disajikan pada Tabel 3 sebagai berikut.

C0	1,2,3,6,11,12,13,14,15,16,17,18	12
C1	4,5,7,8,9,10,19,20	8

Tabel Pengelompokan hasil Iterasi menunjukkan jumlah objek data yang terdistribusi ke dalam dua cluster setelah iterasi dilakukan. Cluster C0 terdiri dari 12 anggota, sementara cluster C1 memiliki 8 anggota. Pengelompokan ini mencerminkan hasil iterasi di mana objek-objek data telah dikelompokkan berdasarkan kedekatannya dengan centroid masing-masing cluster.

Pada tabel diatas terdapat pengelompokan hasil iterasi. iterasi C0 jumlah anggota yaitu 12 anggota. Sedangkan iterasi C1 jumlah anggotanya sebanyak 8 anggota.

3.5. Mengulangi ke tahap 3 hingga nilai centroid yang dihasilkan tetap dan anggota cluster tidak berpindah ke cluster lain

Proses iterasi K-Means clustering dilanjutkan dengan mengulangi langkah-langkah sebelumnya hingga posisi

centroid tidak berubah lagi, dan anggota cluster tetap berada di cluster yang sama. Ketika tidak ada perpindahan anggota cluster dan nilai centroid sudah konvergen, proses clustering dianggap selesai. Stabilitas ini menunjukkan bahwa algoritma telah menemukan pengelompokan data yang optimal sesuai dengan pola yang ada. Pada penelitian ini, proses iterasi berhenti pada iterasi ke 4, karna nilai iterasi 4 sudah tidak berubah lagi maka proses algoritma telah selesai disajikan pada Tabel 4.

Nama	C0	C1	Cluster
Sri Wahyuni	9,83248117	2,333333333	C1
Endang	11,53829413	1,054092553	C1
Irma Khairani	3,363636364	8,856886837	C0
Hotlina	1,363636364	10,76000826	C0
Sahriani	5,049343301	17,08150397	C0
Romauli	4,800137739	7,401201104	C0
Titi Namara	2,566471675	9,545214042	C0
Yulia	2,566471675	9,545214042	C0
Siti Sumanti	5,724385997	17,78888542	C0
Fitria Delmayani	0,363636364	11,72367026	C0
Yuni Hesti	16,84543247	4,772607021	C1
Siti Maryam	13,94943644	1,943650632	C1
Risna Yanti	19,77246604	7,731609003	C1
Masdaniah	11,75296692	0,333333333	C1
Ratna Arianti	7,951298038	4,333333333	C1
Febri Yanti	3,509138012	8,569973421	C0
Yusmawarti	8,599442611	3,48010217	C1
Arianti	8,599442611	3,48010217	C1

Penjelasan tabel diatas menunjukkan bahwa hasil iterasi berhenti pada iterasi ke 4. Bahwa posisi cluster terakhir tidak berubah lagi dan ditemukan pada iterasi ke 4. Maka hasilnya didapat bahwa cluster 1 (C0) untuk kategori jangka pendek. Sedangkan cluster 2 (C1) untuk kategori jangka panjang. Tabel Hasil Iterasi 4 menunjukkan distribusi akhir dari objek data setelah proses iterasi K-Means clustering selesai. Pada iterasi ini, posisi centroid sudah tidak berubah lagi, dan anggota cluster tetap berada di cluster yang sama. Cluster C0 terdiri dari objek-objek yang termasuk dalam kategori jangka pendek, sedangkan cluster C1 mencakup objek-objek dalam kategori jangka panjang, menandakan bahwa proses pengelompokan telah mencapai konvergensi. Setelah di analisis, maka proses algoritma K-Means Clustering dapat disimpulkan pada Tabel 5.

Cluster	Jumlah Anggota	Anggota Kelompok
C0	11	3,4,5,6,7,8,9,10,16,19,20
C1	9	1,2,11,12,13,14,15,17,18

Penjelasan tabel di atas yaitu cluster C0 anggota berjumlah 11. Dimana anggota kelompok yaitu 3,4,5,6,7,8,9,10,16,19,20. Cluster C1 anggota kelompok berjumlah 9 anggota. Dimana anggota kelompok yaitu 1,2,11,12,13,14,15,17,18. Tabel 5 Hasil Analisis K-Means Clustering merangkum jumlah anggota yang terdistribusi ke dalam dua cluster setelah proses clustering selesai. Cluster C0 terdiri dari 11 anggota, yang dikelompokkan berdasarkan karakteristik tertentu, sementara Cluster C1 memiliki 9 anggota.

Hasil ini menunjukkan bahwa data telah berhasil diorganisasikan ke dalam dua kelompok yang berbeda, masing-masing mencerminkan pola atau hubungan yang ditemukan dalam data.

4. Kesimpulan

Penelitian ini menunjukkan bahwa penerapan algoritma K-Means Clustering dalam mengelompokkan rekomendasi alat kontrasepsi di Puskesmas terbukti efektif dalam memberikan rekomendasi yang lebih akurat kepada pasangan usia subur. Penggunaan RapidMiner berhasil mengelompokkan data pasien berdasarkan karakteristik seperti nama, umur dan jumlah anak, yang menghasilkan dua cluster, yaitu cluster C0 sebagai jangka pendek dan C1 sebagai jangka panjang. Hasil analisis menunjukkan bahwa algoritma ini dapat membantu tenaga kesehatan, khususnya bidan dalam meningkatkan kualitas pelayanan keluarga berencana. Penelitian ini tidak hanya memberikan kontribusi terhadap pengembangan metode pengelompokan data, tetapi juga berpotensi meningkatkan efisiensi dan efektivitas dalam pengelolaan layanan kesehatan di Masyarakat.

Daftar Rujukan

- [1] Fakhri, D. A., Defit, S., & Sumijan. (2021). Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering. *Jurnal Informasi Dan Teknologi*, 3, 160–166. <https://doi.org/10.37034/jidt.v3i3.137>
- [2] Wandana, J., Defit, S., & Sumijan. (2020). Klasterisasi Data Rekam Medis Pasien Pengguna Layanan BPJS Kesehatan Menggunakan Metode K-Means. *Jurnal Informasi Dan Teknologi*, 2, 199–125. <https://doi.org/10.37034/jidt.v2i4.73>
- [3] Malik Namus Akbar, F. (2024). Metode KNN (K-Nearest Neighbor) untuk Menentukan Kualitas Air. *TEKNO KOMPAK*, 18(1), 1–13.
- [4] Andema, H., Defit, S., & Yuhandri. (2020). *Jurnal Informatika Ekonomi Bisnis Optimalisasi Penggunaan Lahan Perkebunan Kelapa Hibrida Menggunakan K-Means Clustering*. <https://doi.org/10.37034/infeb.v2i2.23>
- [5] Sari, H. L., & Beti, I. Y. (2023). Penerapan Data Mining Dalam Pengelompokan Buku Yang Dipinjam Menggunakan Algoritma K-Means. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 3(6), 925–933. <https://doi.org/10.30865/klik.v3i6.826>
- [6] Oktarian, S., Defit, S., & Sumijan. (2020). Clustering Students' Interest Determination in School Selection Using the K-Means Clustering Algorithm Method. *Jurnal Informasi Dan Teknologi*, 68–75. <https://doi.org/10.37034/jidt.v2i3.65>
- [7] Irianto, M. R., Maududie, A., & Arifin, F. N. (2022). Implementation of K-Means Clustering Method for Trend Analysis of Thesis Topics (Case Study: Faculty of Computer Science, University of Jember). *Berkala Sainstek*, 10(4), 210. <https://doi.org/10.19184/bst.v10i4.29524>
- [8] Maleke, T. S., Pangkey, M., & Tampongangoy, D. (2022). EFEKTIVITAS PROGRAM KAMPUNG KELUARGA BERENCANA (KB) DI DESA TEMBOAN KECAMATAN MAESAN KABUPATEN MINAHASA SELATAN. In *Jurnal Administrasi Publik JAP: Vol. VIII (Issue 2)*.
- [9] Ula, M., Zulfikri, A., Ulva, A. F., & Rizal, R. A. (2023). Penerapan Machine Learning Clustering K-Means dan Linear Regression Dalam Penentuan Tingkat Resiko Tuberkulosis Paru. *Indonesian Journal of Computer Science*, 12(1), 336–348. <https://doi.org/10.33022/ijcs.v12i1.3162>
- [10] Putra Aryadi, B., & Hendrastuty, N. (2024). PENERAPAN ALGORITMA K-MEANS UNTUK MELAKUKAN KLASTERISASI PADA VARIETAS PADI. In *Jurnal Informatika & Rekayasa Elektronika (Vol. 7, Issue 1)*. <http://ejournal.stmiklombok.ac.id/index.php/jireISSN.2620-6900>
- [11] Pane, P. P., Ramadhan Nasution, Y., & Furqan, Mhd. (2024). Implementasi Data Mining dengan K-Means Clustering untuk Memprediksi Pengadaan Obat. *Journal of Computer System and Informatics (JoSYC)*, 5(2), 286–296.
- [12] Sari, H. L., & Beti, I. Y. (2023). Penerapan Data Mining Dalam Pengelompokan Buku Yang Dipinjam Menggunakan Algoritma K-Means. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 3(6), 925–933. <https://doi.org/10.30865/klik.v3i6.826>
- [13] Purba, W., Sembiring, G. A., Saputra, A., Turnip, T., Jua, B., & Manihuruk, I. (2023). PENERAPAN DATA MINING UNTUK PENGELOLAAN DATA REKAM MEDIS MENGGUNAKAN METODE K-MEANS CLUSTERING PADA RUMAH SAKIT ROYAL PRIMA MEDAN. *Jurnal TEKINKOM*, 6(1), 158–166. <https://doi.org/10.37600/tekinkom.v6i1.857>
- [14] Riana, D., Rahayu, S., Hadiani, S., Frieyadi, F., Hasan, M., Karimah, I. N., & Pratama, R. (2022). Identifikasi Citra Pap Smear RepoMedUNM dengan Menggunakan K-Means Clustering dan GLCM. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 6(1), 1–8. <https://doi.org/10.29207/resti.v6i1.3495>
- [15] Amijoyo, T., & Siti Nurhaliza, M. (2023). Application Of K-Means Clustering Algorithm Method In New Student Admissions. *Jurnal Inovatif: Inovasi Teknologi Informasi Dan Informatika*, 6(1), 92–100. <https://doi.org/10.32832/inovatif>