CLUSTERING KUNJUNGAN PASIEN MENGGUNAKAN ALGORITMA K-MEANS PADA RUMAH SAKIT DI WILAYAH BEKASI

Muhammad Faizal Rizqi, Martanto, Umi Hayati
Teknik Informatika, Manajemen Informatika, STMIK IKMI Cirebon
Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Indonesia
faizalr884@gmail.com

ABSTRAK

Dalam era digital saat ini. Sektor kesehatan terus menghadapi tantangan besar dalam mengelola dan menganalisis data pasien yang terus bertambah. Masalah yang terjadi adalah pengelolaan statistik untuk menentukan jumlah kunjungan dari terendah, menengah sampai tertinggi pada setiap rumah sakit dalam skala yang besar. Pada penelitian ini bertujuan untuk mengelompokkan jumlah kunjungan pasien berdasarkan rumah sakit di wilayah Bekasi dari yang terendah, menengah, dan tertinggi. Apabila pengolahan data dilakukan dengan analisis manual maka menjadi tidak efisien dan menghambat waktu, maka dari itu perlunya menggunakan teknik data mining dengan algoritma yang sesuai. Metode yang diterapkan yaitu menggunakan Algoritma K-Means Clustering, karena kemampuannya dalam mengelompokkan data menjadi kelompok yang saling berdekatan berdasarkan nilai terdekatnya. Melalui implementasi algoritma K-Means ini, penelitian berhasil mengelompokkan berdasarkan kunjungan rumah sakit. Hasil evaluasi menunjukkan bahwa jumlah kelompok optimal adalah tiga dengan nilai *DBI* 0,47. Pada cluster cluster 0 menjadi cluster yang terendah dengan jumlah maksimal rawat jalan sebesar 89.360 orang dan rawat inap sebesar 28.179 orang. Untuk cluster 1 menjadi cluster tersedang, dengan jumlah maksimal rawat jalan sebesar 244.966 orang dan rawat inap sebesar 45.919 orang. Dan untuk cluster 2 menjadi cluster yang tertinggi dengan jumlah maksimal rawat jalan sebesar 501.699 orang dan rawat inap 74.419 orang.

Kata kunci : *Algoritma K-Means, Jumlah Kunjungan Pasien, Clustering, Jumlah Kunjungan, Layanan Kesehatan, Statistik Kunjungan*

1. PENDAHULUAN

Pada era modern saat ini, sektor kesehatan menghadapi tantangan besar dalam mengolah dan menganalisis data pasien yang terus bertambah setiap waktunya. Dibutuhkan pengolahan data yang bersifat sistematis, data mining merupakan salah satu metode pengolahan dalam menganalisis data yang belum terstruktur menjadi terstruktur. Maka dari itu data mining menjadi metode yang paling banyak digunakan oleh setiap lembaga dalam mengolah data salah satunya pada bidang Kesehatan. Pengelolaan data kesehatan adalah tugas penting dalam mengelompokkan jumlah kunjungan pasien berdasarkan karakteristik tertentu untuk membantu dalam pengambilan keputusan yang klinis. Kunjungan pasien setiap tahunnya akan terus bertambah banyak dan menyesuaikan dengan kategori kunjungannya, tentunya ini akan menjadi masalah yang besar apabila data tidak dikelola dengan baik. Masalah tersebut bisa dipecahkan dengan data mining menggunakan metode clustering dan menggunakan algoritma K-Means.

Penggunaan data mining pada metode clustering sangat membantu efektifitas layanan statistik pada data Kesehatan. Pengelolaan yang baik dari data kunjungan pasien dapat membantu rumah sakit dalam perancanaan sumber daya, alokasi anggaran serta pelayanan yang lebih terjamin kepada pasien. Implikasi yang dihasilkan diharapkan dapat memberikan wawasan yang lebih mendalam tentang bagaimana menentukan rasio perbandingan jumlah kunjungan kedalam kategori kunjungan yang berbeda. Disamping itu, penelitian ini dapat memberikan kontribusi pada perkembangan teknologi kesehatan pada setiap lembaga kesehatan di daerah yang belum tersentuh teknologi data mining pada pengolahan data. Dengan berkembangnya teknologi data mining pada setiap daerah tentunya dapat memberikan efektifitas dalam pengolahan data. Sehingga pengolahan data menjadi lebih cepat dan tidak membutuhkan waktu yang lama untuk dikelola. Maka dari itu pentingnya penggunaan Teknik data mining dalam pengelolaan sebuah data untuk dianalisa.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data mining melibatkan pemanfaatan teknik statistik, matematika, kecerdasan buatan, dan machine learning dalam rangka mengungkap informasi berharga yang terkait dari kumpulan data besar. Fokus utama dari data mining adalah mendeteksi, mengungkap, atau mengeksplorasi pengetahuan yang dapat diambil dari data atau informasi yang tersedia [1].

2.2. Clustering

Teknik pengelompokan memiliki dua metode, yaitu hierarchical clustering dan non-hierarchical clustering. Hierarchical clustering adalah metode pengelompokan data yang bekerja dengan mengelompokkan dua data atau lebih yang memiliki kesamaan atau kemiripan, kemudian langkah tersebut diulang untuk objek lain yang memiliki kedekatan,

proses ini berlanjut hingga membentuk suatu struktur pohon di mana terdapat hirarki atau tingkatan yang memperlihatkan hubungan antar objek, mulai dari yang paling mirip hingga yang paling tidak mirip [2].

2.3. Algoritma K-Means

K-Means adalah suatu metode algoritma yang digunakan untuk mengelompokkan data dengan cara memisahkan mereka ke dalam kelompok yang berbeda. Algoritma ini memiliki kemampuan untuk mengurangi jarak antara data dan klusternya [3].

2.4. Penelitian Terdahulu

Penelitian pertama yang dilakukan [4] dengan judul "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport" masalah yang terdapat pada penelitian ini adalah untuk mencari strategi meningkatkan penjualan dan pemasaran di toko olahraga tersebut. Hasil Data Performance menunjukkan bahwa Cluster 0 dengan jumlah 110 item dianggap tidak terjual, sedangkan 21389 item dianggap laris. Dalam hasil rekomendasi penelitian ini, ditemukan pola dan informasi dari penerapan algoritma k-means pada data penjualan, dengan 99 item barang terjual dengan baik dan 23 item barang yang tidak terjual. Informasi ini dapat membantu pemilik toko dalam merancang strategi penjualan dan keputusan pembelian berdasarkan barang yang laku terjual.

Penelitian kedua yang dilakukan [5] dengan judul "Klasterisasi Pasien Rawat Inap Peserta BPJS Berdasarkan Jenis Penyakit Menggunakan Algoritma K-Means". Masalah pada penelitian ini yaitu mengatasi pengelompokkan pasien rawat inap BPJS agar sesuai dengan informasi penyebaran penyakit pasien. Penelitian ini mengungkap bahwa kelompok pertama didominasi oleh pasien dengan penyakit Malignant neoplasm, breast, unspecified (C50.9) dan Non-hodgkin's lymphoma, unspecified type (C85.9). Sementara itu, kelompok kedua didominasi oleh pasien dengan penyakit Fracture of neck of femur, closed (S71.00) dan Dengue haemorrhagic fever (A91). Dari hasil pengelompokan ini, dapat disimpulkan bahwa penggunaan algoritma K-Means berhasil dalam mengklasifikasikan data rekam medis pasien BPJS yang dirawat inap menjadi tiga kelompok

Penelitian ketiga yang dilakukan [6] dengan judul "Klasterisasi Data Penanganan Dan Pelayanan Kesehatan Masyarakat Menggunakan Algoritma K-Means". Masalah yang terdapat pada penelitian ini yaitu mengenai penurunan Tingkat pelayanan Kesehatan Masyarakat. Dalam konteks ini, perlu ditekankan pentingnya optimalisasi rencana pelayanan Kesehatan Masyarakat. Salah satu solusi yang diajukan adalah memanfaatkan teknik pengelolaan data Kesehatan dengan menggunakan konsep data mining. Dengan menerapkan metode K-Means dalam pengelompokan data Kesehatan Masyarakat, penelitian ini bertujuan untuk menghasilkan informasi yang lebih baik tentang kondisi kesehatan masyarakat. Hasil penelitian ini menunjukkan bahwa pendekatan menggunakan algoritma K-Means dapat mengklasifikasikan data dengan lebih cepat dan akurat, sehingga dapat memberikan hasil yang maksimal. Untuk menganalisis permasalahan terkait klasterisasi data penanganan dan pelayanan kesehatan masyarakat, penelitian ini melibatkan observasi dan wawancara langsung dengan sumber-sumber yang berkaitan dengan penanganan kesehatan, layanan kesehatan, dan fasilitas yang disediakan oleh Dinas Kesehatan. Pengujian algoritma K-Means dalam klasterisasi data penanganan dan pelayanan kesehatan masyarakat dilakukan melalui serangkaian langkah, termasuk inisialisasi cluster k, penentuan nilai awal centroid, pengelompokan data, pembaruan centroid, dan penampilan hasil klaster.

Penelitian keempat yang dilakukan [7] dengan judul "Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI". Masalah pada penelitian ini yaitu kekurangan dan kelebihan obat pada jumlah yang tidak terlalu banyak. Pembahasan ini mencakup pengelompokkan obat berdasarkan Tingkat pemakaian, yang dapat digunakan Sebagai panduan atau dasar untuk pengelolaan stok obat. Hasil penelitian mengidentifikasi dua kelompok utama dalam data obat, kelompok pertama memiliki tingkat pemakaian tinggi dan terdiri dari 6 jenis obat dengan jumlah pemakaian mencapai 2046 item, sementara kelompok kedua memiliki tingkat pemakaian rendah dengan 933 jenis obat dan jumlah pemakaian di bawah 2046 item.

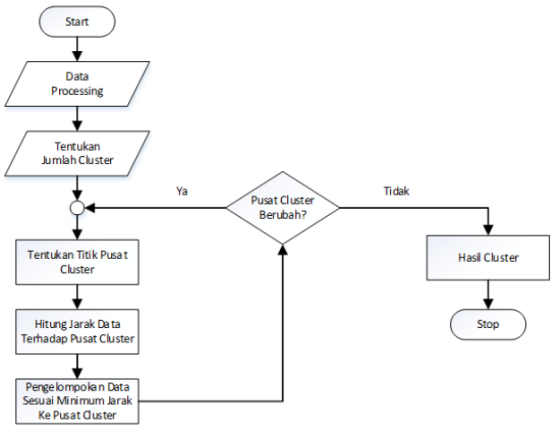
Penelitian kelima yang dilakukan [8] dengan judul "Penerapan Klasterisasi Menggunakan K-Means untuk menentukan Tingkat Kesehatan Bayi dan Balita di Kabupaten Bengkulu Utara". Masalah yang terdapat pada penelitian ini yaitu Tingkat Kesehatan Bayi yang Masih Menjadi Perhatian Dan Kurangnya Pelayanan Kesehatan Dalam Menentukan Tingkat Kesehatan Bayi". Hasil dari proses pengelompokan ini menghasilkan tiga kelompok, yakni tingkat kesehatan tinggi, sedang, dan rendah. Penggunaan klasterisasi ini memudahkan dalam mengidentifikasi tingkat kesehatan bayi dan balita setiap tahunnya, sehingga pelayanan dan tindakan yang diberikan dapat lebih optimal dan disesuaikan dengan kebutuhan. Proses klasterisasi menggunakan algoritma K-Means menghasilkan tiga kelompok tingkat kesehatan bayi dan balita, yaitu kelompok rendah, sedang, dan tinggi. Setiap tahun, anggota dalam kelompok tersebut dapat berubah karena perubahan titik pusat atau centroid.

Penelitian keenam yang dilakukan [9] dengan judul "Data Mining Menggunakan Algoritma K-Means Pengelompokkan Penyebaran Diare di Kota Makassar". Permasalahan pada penelitian ini terkait dengan penyebaran diare. Salah satu Solusi untuk mengatasi permasalahan ini adalah dengan menggunakan teknink data mining, yaitu algoritma k-means clustering untuk mengelompokkan penyebaran diare. Hasil dari penelitian ini mengungkapkan bahwa

terdapat dua kelompok wilayah penyebaran yang dapat diidentifikasi melalui penerapan Algoritma K-Means.

Penelitian ketujuh yang dilakukan [10] dengan judul "Klasterisasi Pasien BPJS dengan Metode K-Means Clustering Guna Menunjang Program Jaminan Kesehatan Nasional di Rumah Sakit Anwar Medika Balong Bendo Sidoarjo. Permasalahan pada penelitian ini yaitu terjadinya penumpukan data didalam database menjadi kurang optimal dalam penggunaannya. Fokus pada penelitian ini adalah mengelompokkan pasien yang memanfaatkan BJPS, dengan metode K-Means Clustering Sebagai pendekatan utama. Hasil penelitian ini menunjukkan bahwa dari 180 pasien yang diteliti, mereka dapat dikelompokkan ke dalam 3 kelompok atau cluster. Terdapat prevalensi yang tinggi dari pasien yang menderita penyakit dari wilayah kecamatan Krian dan Balonbendo, sementara pasien dengan jenis kelamin perempuan lebih mendominasi dalam menderita berbagai penyakit dibandingkan dengan pasien pria.

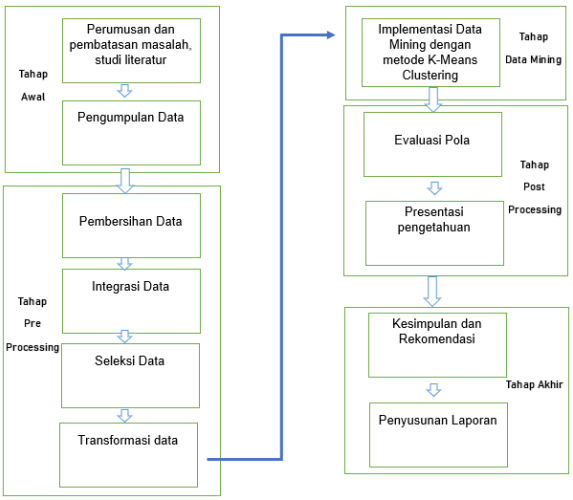
3. METODE PENELITIAN



Gambar 1. Flowchat Algoritma K-Means

Pada metode clustering menggunakan algoritma K-Means terdapat beberapa hal yang harus diperhatikan yaitu pada data processing, dengan menentukan jumlah cluster akan menentukan banyaknya kelompok yang terbentuk. Selanjutnya dengan menentukan titik pusat pada cluster dapat dilakukan secara manual atau acak. Kemudian dilakukan alokasi data untuk menghitung jarak data ke setiap centroid. Menghitung ulang centroid Sebagai rata-rata dari data yang berada dalam cluster. Langkah tersebut dilakukan berulang sampai centroid tidak berubah.

Selanjutnya penelitian ini menggunakan tahapan data mining KDD (Knowledge Discovery In Database) yang terdiri dari tahap awal, preprocessing, tahap data mining, Tahap postprocessing, dan tahap akhir.



Gambar 2. Tahapan KDD

Berikut adalah beberapa tahapan dalam Tahap KDD (Knowledge Discovery In Database):

1. Tahap Awal
   Menentukan perumusan masalah, membuat Batasan masalah dan mencari studi literatur terkait topik penelitian dan mencari data dari berbagai sumber yang relevan dengan penelitian.
2. Tahap Preprocessing
   Melakukan pembersihan data dari missing value, duplikat, serta nilai nol. Melakukan integrasi data dengan menggabungkan data yang terpisah menjadi satu. Seleksi data diperlukan untuk analisis atribut yang akan dipakai. Setelah itu melakukan Transformasi data bertujuan untuk mengubah data dalam sebuah format yang sesuai untuk proses data mining.
3. Tahap Data Mining
   Melakukan implementasi data mining dengan metode K-Means Clustering bertujuan untuk mengolah sebuah data dengan menggunakan sistem pengelompokkan yang akan diolah dengan tools rapidminer.
4. Tahap Postprocessing
   Evaluasi pada hasil yang telah ditemukan pada penerapan data mining K-means Clustering tersebut untuk melihat keakuratan dan ketepatan hasil. Setelah itu melakukan presentasi pengetahuan dengan memberikan informasi yang didapat.
5. Tahap Akhir
   Kesimpulan dan Rekomendasi dari segala aspek hasil penelitian yang telah dilakukan, dan melakukan penyusunan laporan.

4. HASIL DAN PEMBAHASAN

4.1. Tahap Awal (Data)

Pengumpulan data dikumpulkan dengan melakukan pencarian data dengan Teknik sekunder yang didapatkan melalui https://opendata.jabarprov.go.id/ yang berjumlah 191 data.

Table 1. Data Awal

Prov	Kode	Kab / Kot	RumahSakit	R.Jalan	R. Inap	G. Jiwa	Thn
JABAR	3216	KABUPATEN BEKASI	RSU PERMATA KELUARGA JABABEKA	87099	6597	0	2019
JABAR	3216	KABUPATEN BEKASI	RSU KARYA MEDIKA I	67282	7996	1898	2019
JABAR	3216	KABUPATEN BEKASI	RSU DOKTER ADAM THALIB	0	0	0	2019
JABAR	3216	KABUPATEN BEKASI	RSIA MITRA MEDIKA	0	0	0	2019
.	…	…….	……	…...	…...	….	….
JABAR	3275	KOTA BEKASI	RS SATRIA MEDIKA	15379	1958	0	2020

4.2. Tahap PreProcessing

Melakukan pembersihan data pada atribut yang memiliki value 0, menggunakan operator filter example seperti pada gambar 3:

Gambar 3. Operator Filter Example

Bertujuan agar pengolahan data sesuai dengan tujuan penelitian. Pada pembersihan attribut yang memiliki nilai 0 pada rawat jalan dan inap. Untuk gangguan jiwa tidak akan dibersihkan karena nantinya atribut tersebut tidak akan dipakai karna terlalu banyak nilai 0.

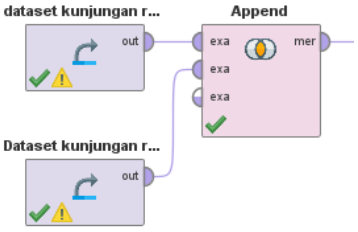
Table 2. Data Sebelum Dibersihkan

Prov	Kode	Kab/Kot	Rumah Sakit	R.Jalan	R. Inap
JABAR	3216	KABUPATEN BEKASI	RSU KARYA MEDIKA I	67282	7996
JABAR	3216	KABUPATEN BEKASI	RSU DOKTER ADAM THALIB	0	0
JABAR	3216	KABUPATEN BEKASI	RSIA MITRA MEDIKA	0	0
.	…	…….	……	…...	…...
JABAR	3275	KOTA BEKASI	RS SATRIA MEDIKA	15379	1958

Dapat dilihat pada tabel bahwa menunjukkan salah satu data yang memiliki nilai 0 tersebut akan dihapus oleh tools rapid miner dengan menggunakan filter example. Maka hasil data yang dibersihkan tertera pada tabel 3:

Table 3. Data Setelah Dibersihkan

Prov	Kode	Kab/Kot	Rumah Sakit	R.Jalan	R. Inap
JABAR	3216	KABUPATEN BEKASI	RSU KARYA MEDIKA I	67282	7996
JABAR	3275	KOTA BEKASI	RS SATRIA MEDIKA	15379	1958



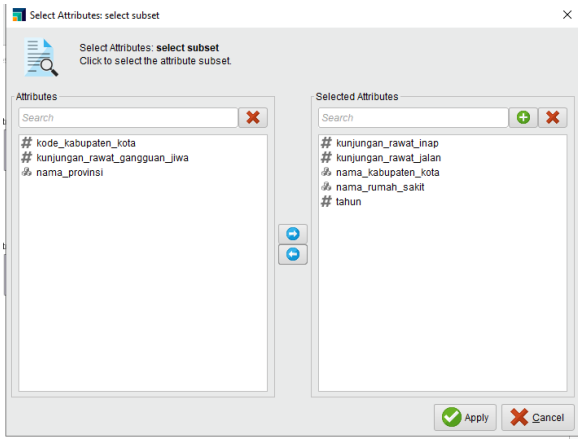
Gambar 4. Operator Append

Melakukan integrasi data yaitu penggabungan pada data untuk menyatukan data menjadi kesatuan. Dalam hal ini data yang akan digabungkan adalah data kunjungan tahun 2019 dan 2020. Menggunakan operator append yang berfungsi untuk menggabungkan data pada rapidminer.
Maka dihasilkan penggabungan data tertera pada tabel 4:

Table 4. Data Setelah Digabung

Prov	Kode	Kab /Kot	RumahSakit	R.Jalan	R. Inap	G. Jiwa	Thn
JABAR	3216	KAB BEKASI	RSUD KABUPATEN BEKASI	83910	12681	3245	2019
JABAR	3275	KOTA BEKASI	RSUD PONDOK GEDE	5199	196	0	2020

Dapat dilihat pada tabel menunjukkan bahwa terdapat dua data yang sudah diintegrasi yaitu kunjungan untuk tahun 2019 dan 2020. Berikutnya yaitu dengan melakukan pemilihan atribut pada data yang akan dipilih dengan menggunakan operator select atribut pada rapidminer.



Gambar 5. Operator Select Attribute

Adapun atribut yang akan dipilih yaitu nama rumah sakit, kabupaten/kota, rawat jalan, rawat inap dan tahun kunjungan. Adapun hasil dari pemilihan atribut tertera pada tabel 5:

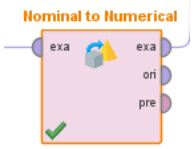
Table 5. Pemilihan Data

Kab/Kot	RumahSakit	R.Jalan	R. Inap	Thn
KAB BEKASI	RSUD KABUPATEN BEKASI	83910	12681	2019
KOTA BEKASI	RSUD PONDOK GEDE	5199	196	2020

Pada Tahap berikutnya melakukan transformasi data untuk mengubah suatu atribut menjadi id agar nantinya pada pengelompokkan dapat mudah dianalisis, dalam hal ini atribut yang dijadikan id ialah nama rumah sakit. Adapun operator yang digunakan yaitu set role.

Gambar 6. Operator set role

Kemudian dengan melakukan pengubahan data nominal menjadi numerical (angka) pada value dari atribut. Dengan menggunakan operator nominal to numerical



Gambar 7. Operator Nominal To Numerical

Maka didapatkan hasil yang tertera pada tabel 6:

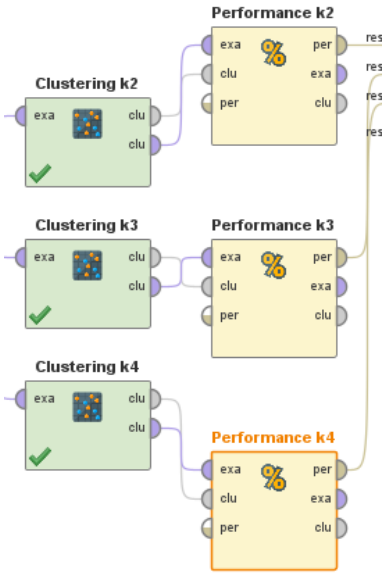
Table 6. Transformasi Data

Kab/ Kot	RumahSakit	R.Jalan	R. Inap	Thn
0	RSUD KABUPATEN BEKASI	83910	12681	2019
1	RSUD PONDOK GEDE	5199	196	2020

Pada angka 0 menunjukkan wilayah kabupaten, sedangkan angka 1 menunjukkan wilayah kota.

4.3. Tahap Data Mining

Pada Tahap ini dilakukan penerapan algoritma dan metode data mining, Adapun algoritma yang digunakan yaitu K-Means dan metode Clustering. Untuk tahap awal menentukan jumlah k pada cluster, hal ini bertujuan untuk melihat jumlah yang efektif dalam metode clustering. Untuk uji cluster yang akurat, pada rapid miner akan membandingkan menjadi tiga kelompok. Pada kelompok satu dengan jumlah k=2, kelompok dua dengan jumlah k=3, dan kelompok tiga dengan jumlah k = 4.



Gambar 8. Operator Clustering

Dapat dilihat pada gambar 8 terdapat 3 dengan nilai k berbeda hal ini dilakukan dengan mencari nilai yang optimal. Dilakukan max runs 10 kali, dengan measure types mixed measures dan dengan jarak Euclidean distance. Dari hasil pengujian untuk nilai k=2 menghasilkan cluster 0 dengan 113 items dan cluster 1 dengan 21 items. Untuk nilai k=3 menghasilkan cluster 0 dengan 109 items dan cluster 1 dengan 2 items dan cluster 2 dengan 23 items sedangkan cluster 2 dengan 23 items. Kemudian dengan nilai k=4 menghasilkan cluster 0 dengan 36 items, cluster 1 dengan 2 items, cluster 2 dengan 18 items dan cluster 3 dengan 78 items.

4.4. Tahap Postprocessing

Dari ketiga pengelompokkan tersebut dilakukan hasil evaluasi performa menggunakan *DBI (Davies Bouldin Index).* Dalam menentukan hasil dari davies bouldin index terdapat beberapa Langkah yang harus diperhatikan yaitu:

1. *Sum Of Square Within Cluster(SSW)*
   Digunakan untuk mencari matriks dalam sebuah cluster ke-i pada persamaan berikut:

$$SSW = \frac{1}{mi}\sum_{j=1}^{mi} d(xj, ci)$$

2. *Sum Of Square Between Cluster (SSB)*
   Persamaan dalam mengenali fungsi pemisah antara cluster

$$SSBi, j = d(ci, cj)$$

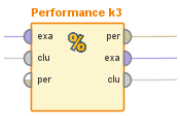
3. *Ratio*
   Membandingkan nilai cluster i dan cluster j

$$RIJ = \frac{SSWi + SSWj}{SSBij}$$

4. *Davies Bouldin Index (DBI)*
   Setelah mendapatkan nilai pada rasio, mencari nilai *DBI*

$$DBI = \frac{1}{k}\sum_{i=1}^{k} max_{i \neq j}\,(R_{i,j})$$

Dalam melakukan uji hasil DBI digunakan operator performa yang sesuai, adapun operator yang digunakan yaitu menggunakan operator cluster distance performance.



Gambar 9. Operator Cluster Distance Performance

Operator ini berfungsi untuk menghitung jarak antara dua cluster. Jarak ini dapat digunakan untuk mengukur seberapa mirip atau berbeda dua cluster tersebut. Maka didapatkan hasil pada masing-maing pengelompokkan cluster didapatkan hasil yang tertera pada tabel 7:

Table 7. Hasil Evaluasi DBI

Nilai (K)	DBI (Davies Bouldin Index)
2	0.548
3	**0.473**
4	0.493

Dari hasil yang tertera pada tabel, didapatkan nilai k yang optimal dengan nilai k=3 yang mendapatkan hasil *DBI (Davies Bouldin Index)* 0.473 karena memiliki nilai yang rendah diantara cluster yang lain.

4.5. Tahap Akhir (Pembahasan)

Pada Tahap akhir dilakukan pembahasan analisis pada hasil penelitian yang telah didapatkan sesuai dengan tujuan dari penelitian yang ditentukan.

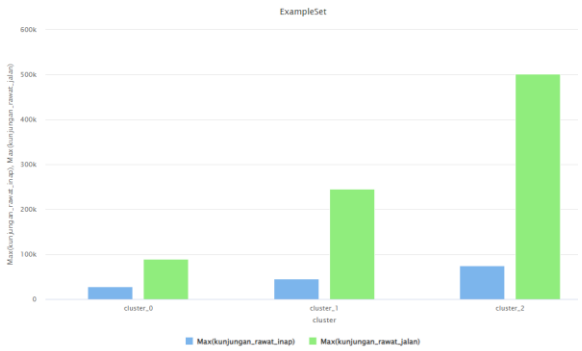
a) Penentuan Jumlah Cluster Yang Optimal
   Dalam menentukan jumlah cluster yang optimal pada metode clustering, tentunya banyak Teknik yang dapat diterapkan untuk penentuan jumlah cluster, dalam hal ini peneliti akan memilih penentuan jarak dengan Teknik Euclidean Distance yang akan berfokus dengan jarak seberapa dekat dan jauh antar kelompok. Dalam penelitian ini membagi menjadi tiga pengujian dengan masing-masing nilai k berbeda. Untuk pengujian pertama dengan nilai k=2, dan pengujian kedua dengan nilai k=3, dan pengujian ketiga dengan nilai k=4. Hal itu bertujuan agar dapat mengetahui dari setiap nilai k yang ditentukan memiliki hasil yang akurat sesuai dengan performa pada masing-masing pengujian. Adapun untuk mengetahui akurat atau tidaknya tentunya disesuaikan dengan metode pengujian pada data mining itu sendiri, dalam hal ini peneliti memilih pengujian peforma nilai k dengan menggunakan perhitungan peforma DBI (Davies Bouldin Index) yang dimana nilai k yang memiliki hasil paling rendah maka dinyatakan sebagai kelompok yang optimal.

b) Evaluasi Performa Menggunakan DBI
   Sesuai dengan hasil pada nilai DBI (Davies Bouldin Index) pada nilai k=3 menjadi nilai yang optimal karena memiliki nilai akurasi performa sebesar 0.473 maka dari itu penentuan nilai k=3 akan dipilih untuk dianalisis data tersebut. Sesuai dengan hasil pengelompokkan dengan nilai k yang optimal menghasilkan menjadi beberapa cluster yang tertera pada tabel 8:

Table 8. Hasil Cluster K=3

Cluster	Jumlah
Cluster 0	109 items
Cluster 1	23 items
Cluster 2	2 items

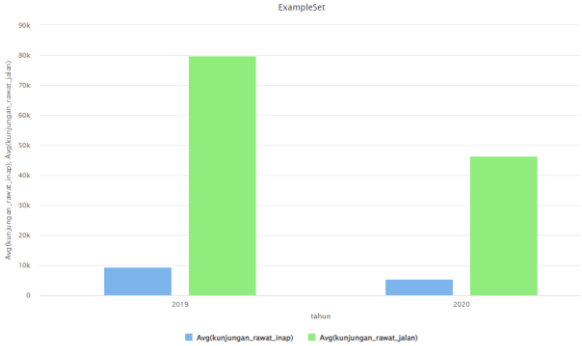
c) Analisis Jumlah Kunjungan Tertinggi dan Terendah
   Dari hasil pengelompokkan pada setiap cluster apabila divisualisasikan dalam bentuk grafik tertera pada gambar 10:



Gambar 10. Analisis Tingkat Kunjungan

Dapat dilihat pada gambar 10, menunjukkan bahwa cluster 0 dengan jumlah maksimal kunjungan rawat jalan dan inap terendah, sedangkan untuk cluster 1 dengan jumlah maksimal kunjungan rawat jalan dan inap tersedang, dan untuk cluster 2 menjadi cluster dengan jumlah maksimal kunjungan rawat jalan dan inap tertinggi. Apabila ingin mengetahui rasio tahun kunjungan dari tahun 2019 dan 2020 dapat dilihat pada gambar 11:

Dapat dilihat pada gambarc11, menunjukkan bahwasannya rata-rata jumlah kunjungan rawat jalan dan rawat inap tahun 2019 menjadi yang paling tinggi dibanding tahun 2020 yang mengalami penurunan jumlah kunjungannya. Pada tahun 2019 jumlah kunjungan rawat jalan menunjukkan rata-rata sebesar 79,795 dan rawat inap sebesar 9,371. Sedangkan rata-rata kunjungan rawat jalan pada tahun 2020 sebesar 46,404 dan rawat inap sebesar 5,394.



Gambar 11. Rasio Tahun Kunjungan

Hal tersebut bisa disebabkan pada tahun 2020 yang merupakan tahun pandemi yang dimana kebijakan pemerintah untuk melakukan PPKM, sehingga enggannya Masyarakat untuk melakukan kunjungan Kesehatan, berdampak pada menurunnya jumlah kunjungan kefasilitas pelayanan Kesehatan Tingkat lanjut yaitu rumah sakit.

Berikut adalah tabel 9 pengelompokkan hasil cluster 0 dengan keterangan Rendah

Table 9 Hasil Cluster 0 Tingkat Kunjungan Rendah

No	RS	Cluster	Kab/Kot	R.Jalan	R.Inap	Thn	Ket
1	RSUD KABUPATEN BEKASI	cluster_0	0	83910	12681	2019	Rendah
2	RSU PERMATA BUNDA	cluster_0	0	6759	964	2019	Rendah
3	RSUD CABANG BUNGIN	cluster_0	0	1219	26	2019	Rendah
…	…	…	…	…	…	…	…
109	RS SATRIA MEDIKA	cluster_0	1	15379	1958	2020	Rendah

Berikut adalah tabel 10 pengelompokkan hasil cluster 1 dengan keterangan Sedang

Table 10. Hasil Cluster 1 Tingkat Kunjungan Sedang

No	RS	Cluster	Kab/Kot	R.Jalan	R.Inap	Thn	Ket
1	RSU CIBITUNG MEDIKA	cluster_1	0	158831	20826	2019	Sedang
2	RSU ANNISA	cluster_1	0	193732	17942	2019	Sedang
3	RSU MITRA KELUARGA CIKARANG	cluster_1	0	124467	8690	2019	Sedang
…	…	…	…	…	…	…	…
23	RS ANNA MEDIKA	cluster_1	1	124934	9330	2020	Sedang

Berikut adalah tabel 11 pengelompokkan hasil cluster 2 dengan keterangan Tinggi

Table 11 Hasil Cluster 2 Tingkat Kunjungan Tinggi

No	RS	Cluster	Kab/Kot	R.Jalan	R.Inap	Thn	Ket
1	RS UMUM H DR. CHASBULLAH ABDULMADJID	cluster_2	1	323363	46364	2019	Tinggi
2	RS UMUM HERMINA BEKASI	cluster_2	1	501699	74419	2019	Tinggi

5. KESIMPULAN DAN SARAN

Berdasarkan dari hasil penelitian yang sudah diuaikan sebelumnya, maka peneliti menyimpulkan bahwasannya Penentuan jumlah cluster optimal dapat ditentukan menggunakan tiga kali pengujian nilai k berbeda dengan menggunakan jarak Euclidean distance dan uji nilai DBI. Evaluasi yang digunakan menggunakan DBI (Davies Bouldin Index) dengan

mendapatkan pengujian nilai k, untuk nilai k=3 menjadi nilai k yang paling rendah berdasarkan hasil performa pada DBI dengan nilai akurasi sebesar 0.473 dan dinyatakan sebagai jumlah cluster yang optimal. Menunjukkan bahwa cluster 0 menjadi cluster yang terendah dengan jumlah maksimal rawat jalan sebesar 89.360 orang dan rawat inap sebesar 28.179 orang. Untuk cluster 1 menjadi cluster tersedang, dengan jumlah maksimal rawat jalan sebesar 244.966 orang dan rawat inap sebesar 45.919 orang. Dan untuk cluster 2 menjadi cluster yang tertinggi dengan jumlah maksimal rawat jalan sebesar 501.699 orang dan rawat inap 74.419 orang. Peneliti menyarankan bahwa kepada penelitian berikutnya yang akan mengambil penelitian serupa, bahwasannya pada jumlah atribut agar bisa ditambah, karena untuk melengkapi tingkat kelengkapan informasi pada data contohnya, menambah atribut fasilitas pada rumah sakit, serta menambah data untuk tahun terbaru agar rasio kunjungan bisa dianalisis lebih baik. Dan untuk penentuan jumlah k agar bisa disesuaikan dengan tujuan dari penelitian.

DAFTAR PUSTAKA

[1] D. F. Pasaribu, I. S. Damanik, E. Irawan, Suhada, and H. S. Tambunan, "Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat," *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 2, no. 1, pp. 11–20, 2021, doi: 10.37148/bios.v2i1.17.

[2] A. M. G. S. Et, "Algoritma K-Means Dalam Mengelompokkan Desa / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Berdasarkan Provinsi," *Semin. Nas. Teknol. Komput. Sains SAINTEKS 2019*, pp. 754–761, 2018.

[3] P. Alkhairi and A. P. Windarto, "Penerapan K-Means Cluster pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara," *Semin. Nas. Teknol. Komput. Sains*, pp. 762–767, 2019.

[4] A. Nugraha *et al.*, "PENERAPAN DATA MINING METODE K-MEANS CLUSTERING UNTUK," vol. 6, no. 2, pp. 849–855, 2022.

[5] Y. Saputra Sy, "Klasterisasi Pasien Rawat Inap Peserta BPJS Berdasarkan Jenis Penyakit Menggunakan Algoritma K-Means," *J. Sistim Inf. dan Teknol.*, vol. 5, pp. 33–37, 2022, doi: 10.37034/jsisfotek.v5i2.162.

[6] S. U. Tarigan, S. Saniman, and M. Yetri, "Klasterisasi Data Penanganan Dan Pelayanan Kesehatan Masyarakat Menggunakan Algoritma K-Means," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 3, p. 193, 2022, doi: 10.53513/jursi.v1i3.5223.

[7] M. R. Nugroho, I. E. Hendrawan, and P. P. Purwantoro, "Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI," *Nuansa Inform.*, vol. 16, no. 1, pp. 125–133, 2022, doi: 10.25134/nuansa.v16i1.5294.

[8] D. T. Saputro and W. P. Sucihermayanti, "Penerapan Klasterisasi Menggunakan K-Means untuk Menentukan Tingkat Kesehatan Bayi dan Balita di Kabupaten Bengkulu Utara," vol. 12, pp. 146–155, 2021.

[9] S. Hasyrif, Rismayani, and S. Asrul, "Data Mining Menggunakan Algoritma K-Means Pengelompokan Penyebaran Diare Di Kota Makassar," *SISITI Semin. Ilm. Sist. Inf. dan Teknol. Inf.*, vol. VIII, no. 1, pp. 73–82, 2019.

[10] A. Ali, L. Masyfufah, S. Yayasan Rumah Sakit DrSoetomo, and K. Kunci, "Klasterisasi Pasien Bpjs Dengan Metode K-Means Clustering Guna Menunjang Program

[11] Jaminan Kesehatan Nasional Di Rumah Sakit Anwar Medika Balong Bendo Sidoarjo," *J. Wiyata*, pp. 8–22, 2020.