

Klasterisasi Penyakit HIV/AIDS di Jawa Barat Menggunakan Algoritma K-Means Clustering

Wina Rosida, Yudhistira Arie Wijaya

Program Studi Manajemen Informatika, STMIK IKMI Cirebon, Cirebon, Indonesia

INFORMASI ARTIKEL

Diterima Redaksi: 19 Februari 2023
Revisi Akhir: 26 Maret 2023
Diterbitkan Online: 29 Maret 2023

KATA KUNCI

Klasterisasi; HIV/AIDS; Algoritma K-Means

KORESPONDENSI

Phone: 082119801840
E-mail: sintawinarosida@gmail.com

ABSTRAK

HIV (*Human Immunodeficiency Virus*) adalah virus yang menyerang system kekebalan tubuh yang dapat melemahkan kemampuan tubuh melawan infeksi dan penyakit. Pada tahun 2017 angka tertinggi infeksi HIV/AIDS di Indonesia yaitu di DKI Jakarta (46.378 orang), diikuti Jawa Timur (33.043 orang), Papua (25.586 orang), Jawa Barat (24.650 orang), dan Jawa Tengah (18.038 orang). Secara nasional, Jawa Barat masih termasuk kedalam empat provinsi dengan kasus HIV/AIDS tertinggi. Hal ini menunjukkan bahwa HIV/AIDS sudah menjadi ancaman bagi masyarakat luas karena selain mengancam nyawa penderitanya, penyakit ini berisiko terhadap penularan yang akan semakin meningkat. Meningkatnya kasus HIV/AIDS dapat menjadi masalah dalam kejiwaan penderitanya, karena penyakit ini dapat menimbulkan dampak negative berupa masalah fisik, psikologis, sosial dan spiritual yang menyebabkan ODHA (Orang dengan HIV AIDS) menjalani kehidupan dengan penuh tekanan. Dalam penelitian ini dilakukan perhitungan menggunakan algoritma K-Means dengan Optimize Parameters Grid pada data kasus penyebaran HIV/AIDS pada tahun 2019-2021 bersumber dari Dinas Kesehatan Provinsi Jawa Barat. K-Means merupakan salah satu algoritma dalam data mining yang bisa digunakan untuk pengelompokan/clustering suatu data. Adapun data yang digunakan dalam penelitian ini sebanyak 971 record. Tujuan penelitian ini dilakukan untuk mengetahui cluster dari penyebaran penyakit HIV/AIDS sebagai salah satu upaya untuk membantu pemerintah dalam menekan angka kasus HIV/AIDS di provinsi Jawa Barat. Hasil dari penelitian ini adalah melakukan perbandingan DBI dengan metode K-Means dari k-2 sampai k-20 yang terdapat pada tabel diatas, dapat dilihat bahwa cluster yang mendekati 0 yaitu k-2, fengan nilai DBI sebesar 0,414. Karena nilai k-2 merupakan nilai terkecil dibandingkan k lainnya, maka dapat disimpulkan bahwa k-2 dengan nilai 0,414 yang paling mendekati 0 merupakan hasil *cluster* terbaik.

PENDAHULUAN

Kasus HIV/AIDS setiap tahunnya terus meningkat, pada tahun 2017 dilaporkan jumlah kumulatif penderita HIV/AIDS di Indonesia berjumlah 242.699 orang [1]. Prevalensi kasus HIV/AIDS masih menjadi masalah Kesehatan di Indonesia. Secara nasional, Jawa Barat masih termasuk ke dalam lima provinsi dengn kasus HIV/AIDS tertinggi [2]. HIV (*Human Immunodeficiency Virus*) adalah virus yang menyerang system kekebalan tubuh yang dapat melemahkan kemampuan tubuh untuk melawan infeksi dan penyakit. Sedangkan AIDS adalah kondisi dimana HIV sudah pada tahap infeksi akhir. Pada tahap ini kemampuan tubuh untuk melawan infeksi yang masuk sangat rendah sehingga sangat rentan terhadap berbagai infeksi [3].

Metode yang digunakan untuk klasterisasi kasus penyebaran HIV/AIDS di Jawa Barat yaitu de ngan menggunakan algoritma *K-Means*. Algoritma *K-Means* merupakan salah satu algoritma yang banyak digunakan dalam pengelompokan karena kesederhanaannya dan efisiensi secara partisi yang memisahkan data ke dalam kelompok yang berbeda-beda [4]. Algoritma *K-Means* merupakan algoritma yang mampu meminimalkan jarak antara data ke clusternya [3]. *Clustering*

merupakan metode yang cukup populer dan paling sering digunakan dalam pengolahan data karena kesederhanaannya dan sangat efektif dalam mengelompokkan data dengan ukuran besar. *K-Means* merupakan salah satu metode clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster [5].

Berdasarkan penelitian sebelumnya pada Jurnal Penerapan Sistem Informasi yang dilakukan oleh Rut Indra Lita Sinaga, Widodo Saputra, dan Hendry Qurniawan pada tahun 2021 dengan judul “Pengelompokan Jumlah Kasus Penyakit AIDS Berdasarkan Provinsi Menggunakan Metode K-Means”. AIDS merupakan salah satu penyakit infeksi peringkat atas yang dapat menyebabkan kematian. Kasus penderita yang meninggal akibat HIV/AIDS juga menjadi faktor penentu kenaikan potensi tersebut. Metode yang digunakan dalam penelitian ini yaitu dengan menggunakan algoritma K-Means. Algoritma K-Means merupakan suatu algoritma yang digunakan dalam pengelompokan secara partisi yang memisahkan data ke dalam kelompok yang berbeda-beda. Tujuan penelitian ini untuk mengetahui hasil akhir perhitungan K-Means *Clustering* dalam mengelompokkan jumlah kasus penyakit AIDS berdasarkan Provinsi. Dengan adanya hasil penelitian tersebut diharapkan kepada pemerintah pusat lebih memperhatikan provinsi yang tinggi dengan kasus penderita AIDS untuk mendapat tindak lanjut [3].

Berdasarkan penelitian sebelumnya [2] yang dilakukan terhadap enam partisipan ODHA di Komunitas Rumah Cemara Kota Bandung terdapat enam tema strategi koping yaitu: 1) Pasrah pada Tuhan 2) Tidak ingin memikirkan penyakit 3) Pengalihan masalah, 4) Berusaha untuk bangkit 5) Upaya mengatasi masalah 6) Menyembunyikan status Kesehatan. Berdasarkan ke enam tema yang ditemukan, partisipan ODHA cenderung melakukan emotional focused coping. Adapun tema Menyembunyikan status kesehatan yang tidak sesuai dengan teori koping manapun namun didukung oleh teori adaptasi dan penelitian yang menyatakan ODHA menyembunyikan status kesehatan. Sehingga diharapkan perawat dapat menjalankan strategi intervensi dengan memberikan konseling maupun edukasi strategi koping yang adaptif kepada penderita dan keluarganya dengan cara membantu memberikan alternatif penyelesaian masalah serta melatih ketrampilan ODHA agar memiliki kemampuan dalam menyelesaikan masalah secara konstruktif.

Tujuan penelitian ini dilakukan untuk mengetahui cluster dari penyebaran penyakit HIV/AIDS sebagai salah satu upaya untuk membantu pemerintah dalam menekan angka kasus HIV/AIDS di provinsi Jawa Barat. Adapun yang menjadi alasan dilakukannya penelitian ini adalah untuk mengetahui cluster kasus penyebaran HIV/AIDS yang ada di Jawa Barat berdasarkan parameter umur sehingga menjadi salah satu upaya untuk membantu pemerintah dalam menekan angka kasus HIV/AIDS di Jawa Barat.

TINJAUAN PUSTAKA

Data Mining

Data Mining adalah proses untuk menemukan kolerasi, pola, dan tren baru yang bermakna dengan memilah-milah data dalam jumlah besar yang disimpan di dalam repositori, menggunakan teknologi pengenalan pola serta teknik statistik dan matematika [6]. Salah satu metode yang terjadi di dalam data mining adalah *clustering*. Clustering merupakan salah satu metode Analisa data yang dapat digunakan memecahkan masalah dalam suatu pengelompokan data. Data mining sering juga disebut *knowledge Discovery in Database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [7]. Data Mining merupakan inti dari proses *knowledge Discovery in Database* (KDD) yang melibatkan algoritma untuk mengeksplorasi data, mengembangkan model dan menemukan pola yang sebelumnya tidak diketahui juga dikenal sebagai *patteran recognition* yang digunakan untuk menemukan pola tersembunyi dari data yang diolah [8].

Knowledge in Database (KDD)

Knowledge in Database (KDD) adalah metode yang digunakan untuk dapat memperoleh pengetahuan yang berasal dari database yang ada. Hasil pengetahuan yang diperoleh dapat dimanfaatkan untuk basis pengetahuan (knowledge base) yang digunakan dalam keperluan mengambil keputusan [4]. *Knowledge in Database* (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Data Mining merupakan inti dari proses *knowledge Discovery in Database* (KDD) yang melibatkan algoritma untuk mengeksplorasi data, mengembangkan model dan menemukan pola yang sebelumnya tidak diketahui juga dikenal sebagai *patteran recognition* yang digunakan untuk menemukan pola tersembunyi dari data yang diolah [8].

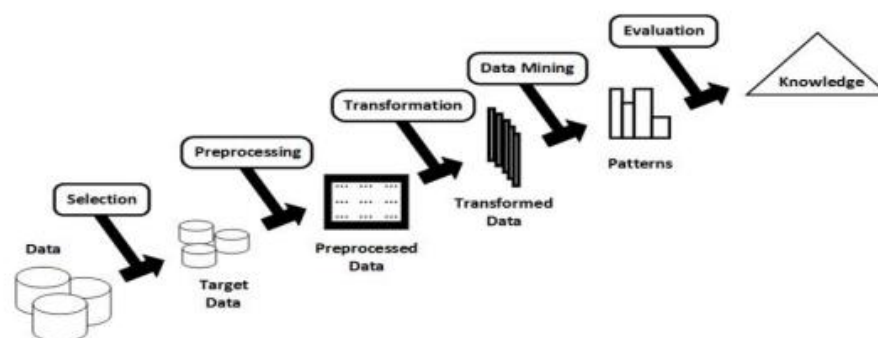
Clustering

Clustering merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar daripada kesamaan data tersebut dengan data dalam kelompok lain. Pada dasarnya clustering merupakan metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (similarity) antara satu data dengan data yang lain[9].

Algoritma K-means Clustering

Algoritma K-Means merupakan algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah k cluster yang sudah ditetapkan di awal[4]. Algoritma K-Means memiliki ketelitian yang cukup tinggi terhadap ukuran objek, sehingga algoritma ini relatif lebih terukur dan efisien untuk pengolahan objek dalam jumlah besar. Selain itu, algoritma K-Means ini tidak terpengaruh terhadap urutan objek. Salah satu tahapan penting dalam menerapkan K-Means Cluster adalah menentukan centroid, banyaknya cluster dan jarak centroid[8].

METODOLOGI



Gambar 1. Tahapan Proses KDD

Teknik Pengumpulan Data

Pada pengumpulan data yang berkaitan dengan penelitian, penulis menggunakan Teknik pengumpulan data skunder untuk mendapatkan hasil yang akurat. Data penelitian terbagi menjadi dua yaitu data primer dan data sekunder. Data primer adalah data yang diperoleh peneliti secara langsung (dari tangan pertama), sementara data sekunder adalah data yang diperoleh peneliti dari sumber yang sudah ada. Contoh data primer adalah data yang diperoleh dari responden melalui kuesioner, kelompok fokus, dan panel, atau juga data hasil wawancara peneliti dengan narasumber. Contoh data sekunder misalnya catatan atau dokumentasi perusahaan berupa absensi, gaji, laporan keuangan publikasi perusahaan, laporan pemerintah, data yang diperoleh dari majalah, dan lain sebagainya.

Tabel 1. Dataset kasus HIV/AIDS di Jawa Barat

Id	Nama Kabupaten Kota	Kelompok Umur	Jenis Keamin	Jumlah Kasus	Tahun
1	KABUPATEN BOGOR	0-4	LAKI-LAKI	4	2019
2	KABUPATEN BANDUNG	0-4	PEREMPUAN	1	2019
3	KABUPATEN SUKABUMI	5-14	LAKI-LAKI	1	2019
4	KABUPATEN SUKABUMI	5-14	PEREMPUAN	0	2019
5	KABUPATEN CIANJUR	15-19	LAKI-LAKI	0	2019
6	KABUPATEN CIANJUR	15-19	PEREMPUAN	0	2019
7	KABUPATEN INDRAMAYU	20-24	LAKI-LAKI	12	2020

8	KABUPATEN INDRAMAYU	20-24	PEREMPUAN	15	2020
9	KABUPATEN PURWAKARTA	25-49	LAKI-LAKI	123	2020
10	KABUPATEN PURWAKARTA	25-49	PEREMPUAN	49	2020
...					
970	KOTA BANJAR	≥50	LAKI-LAKI	2	2021
971	KOTA BANJAR	≥50	PEREMPUAN	1	2021

Adapun data yang digunakan dalam penelitian ini sebanyak 971 baris dan terdapat 6 kolom yaitu id, nama kabupaten kota, kelompok umur, jenis kelamin, jumlah kasus, tahun. (jelaskan kasus yang dari data dibawah). yang bersumber dari Dinas Kesehatan Provinsi Jawa Barat.

Tahap Perancangan

Dalam tahap perancangan ini menggunakan tahapan Knowledge Discovery in Database (KDD), Adapun tahapan KDD adalah sebagai berikut:

1. Data Selection

Pada tahap ini dilakukan proses seleksi atribut mana yang akan digunakan dalam proses data mining. Dari 10 atribut yaitu Id, Nama Kabupaten Kota, Kelompok Umur, Jenis Kelamin, Jumlah Kasus, Tahun, Satuan, Kode Provinsi, Nama Provinsi, Kode Kabupaten Kota. Atribut tersebut akan diseleksi menjadi 6 atribut yaitu Id, Nama Kabupaten Kota, Kelompok Umur, Jenis Kelamin, Jumlah Kasus, Tahun.

2. Preprocessing

Preprocessing merupakan proses pembersihan data dari missing value atau noise yaitu data yang tidak relevan atau tidak konsisten. Pada tahap preprocessing ini, data yang diperoleh akan dibersihkan dari kesalahan, missing, atau ketidak lengkapan kriteria data.

3. Data Transformation

Data Transformation atau tranformasi data merupakan proses perubahan tipe data menjadi tipe data numerik, agar data dapat diolah menggunakan algoritma K-Means.

4. Data Mining

Tahapan data mining ini dilakukan dengan cara menerapkan algoritma atau pencarian pengetahuan. Pada penelitian ini, algoritma yang digunakan adalah algoritma k-means clustering.

5. Interpretation/Evaluation

Pada tahap interpretasi atau evaluasi, dilakukan dengan menganalisa hasil eksperimen yang telah dilakukan.

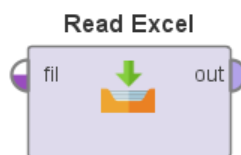
HASIL DAN PEMBAHASAN

Pengelompokan Dataset menggunakan Algoritma K-Means Clustering

Penerapan algoritma k-means clustering ini dirancang dengan menggunakan tool Rapidminer. Pada penelitian ini, menggunakan tahapan Knowledge Discovery in Database (KDD). Penerapan KDD dalam penelitian ini yaitu sebagai berikut:

Data Selection

Read Excel, pada aplikasi RapidMiner fitur Read Excel berfungsi sebagai pembaca file excel. Operator ini digunakan untuk mengimport atau memasukan data excel yang berada di computer pengguna ke dalam proses rapidminer. Read Excel merupakan operator yang paling dasar digunakan sebelum memulai sebuah proses. Operator ini dapat digunakan untuk memuat data dari spreadsheet Microsoft Excel. Dataset yang digunakan adalah dataset kasus HIV/AIDS di Jawa Barat.



Gambar 2. Operator Read Excel

Parameter pada operator Read Excel menggubakan parameter default. Dari hasil pembacaan operator *Read Excel* didapat informasi sebagai berikut.

Tabel 2. Statik Dataset File Pertama

No	Uraian	Keterangan
1.	<i>Record</i>	971
2.	<i>Special Attributes</i>	0
3.	<i>Regular Attributes</i>	10
4.	<i>Attributes:</i>	
	Kode Provinsi	<i>Integer, missing 0</i>
	Nama Provinsi	<i>Polynomial, missing 0</i>
	Kode Kabupaten Kota	<i>Integer, missing 0</i>
	Nama Kabupaten Kota	<i>Polynomial, missing 0</i>
	Kelompok Umur	<i>Polynomial, missing 0</i>
	Jenis Kelamin	<i>Polynomial, missing 0</i>
	Jumlah Kasus	<i>Integer, missing 0</i>
	Satuan	<i>Polynomial, missing 0</i>
	Tahun	<i>Integer, missing 0</i>

Set Role, fitur ini berfungsi untuk membedakan baris penamaan atribut koordinat dan prediksi posisi yang akan di masukan kedalam kategori 'label'. Agar pada saat pengkategorian data 'label' tidak ikut serta terhitung dan merubah hasil.



Gambar 3. Operator Set Role

Parameter pada operator Set Role yang digunakan tampak pada tabel dibawah ini.

Tabel 2. Statik Dataset File Kedua

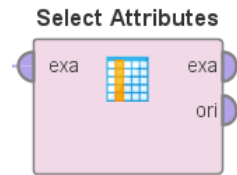
No	Parameter	Isi
1.	<i>Attribute Filter Type</i>	<i>Subset</i>
2.	<i>Selected Attributes</i>	id, jenis kelmin, jumlah kasus, kelompok umur, nama kabupaten kota, tahun

Tabel 3. Statik Dataset File Ketiga

No	Uraian	Keterangan
1.	<i>Record</i>	971
2.	<i>Special Attributes</i>	1
3.	<i>Regular Attributes</i>	9
4.	<i>Attributes :</i>	
	Kode Provinsi	<i>Integer, missing 0</i>
	Nama Provinsi	<i>Polynomial, missing 0</i>
	Kode Kabupaten Kota	<i>Integer, missing 0</i>
	Nama Kabupaten Kota	<i>Polynomial, missing 0</i>
	Kelompok Umur	<i>Polynomial, missing 0</i>
	Jenis Kelamin	<i>Polynomial, missing 0</i>
	Jumlah Kasus	<i>Integer, missing 0</i>

Satuan	<i>Polynomial, missing 0</i>
Tahun	<i>Integer, missing 0</i>

Select Attributes, operator ini digunakan untuk memfilter data apa saja yang akan digunakan dalam pemrosesan data. Atribut dalam dataset kasus HIV/AIDS di filter dari 10 atribut menjadi 6 atribut yang digunakan. Atribut yang digunakan yaitu id, jenis kelamin, jumlah kasus, kelompok umur, nama kabupaten kota, tahun.



Gambar 4. Fitur Select Attributes

Parameter pada operator *Select Attributes* yang digunakan tampak pada tabel dibawah ini.

Tabel 4. Statistik Dataset File Keempat

No	Parameter	Isi
1.	<i>Attribute Filter Type</i>	<i>Subset</i>
2.	<i>Selected Attributes</i>	id, jenis kelamin, jumlah kasus, kelompok umur, nama kabupaten kota, tahun

Dari hasil pembacaan operator *Select Attribute* didapat informasi sebagai berikut.

Tabel 5. Statistik Dataset File Kelima

No	Uraian	Keterangan
1.	<i>Record</i>	971
2.	<i>Special Attributes</i>	1
3.	<i>Regular Attributes</i>	5
4.	<i>Attributes :</i>	
	Id	<i>Integer, missing 0</i>
	Nama Kabupaten Kota	<i>Polynomial, missing 0</i>
	Kelompok Umur	<i>Polynomial, missing 0</i>
	Jenis Kelamin	<i>Polynomial, missing 0</i>
	Jumlah Kasus	<i>Integer, missing 0</i>
	Tahun	<i>Integer, missing 0</i>

Preprocessing

Dikarenakan pada dataset kasus HIV/AIDS tidak ditemukan *missing value* atau data yang tidak memiliki nilai maka *preprocessing* tidak dilakukan.

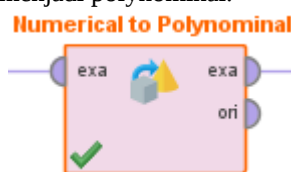
id	Integer	0	Min 1	Max 971	Average 486
nama_kabupaten_kota	Polynomial	0	Least KOTA TASIKMALAYA (3...	Most KABUPATEN BANDUNG...	Values KABUPATEN BANDUNG (36), KABUPATEN BANDUNG BARAT (36), ...[25 more]
kelompok_umur	Polynomial	0	Least ≥50 (161)	Most 0-4 (162)	Values 0-4 (162), 15-19 (162), ...[4 more]
jenis_kelamin	Polynomial	0	Least PEREMPUAN (485)	Most LAKI-LAKI (486)	Values LAKI-LAKI (486), PEREMPUAN (485)
jumlah_kasus	Integer	0	Min 0	Max 254	Average 13.868
tahun	Integer	0	Min 2019	Max 2021	Average 2019.999

Gambar 5. Hasil Statistik

Setelah dilihat dari hasil statistik, tidak ditemukan adanya *missing value*. Maka proses dapat dilanjutkan ke tahap selanjutnya.

Data Transformation

Nominal to Polynominal, digunakan untuk mengubah tipe atribut yang bersifat non-numerik menjadi tipe polynominal. Dalam dataset kasus HIV/AIDS atribut tahun menjadi polynominal.



Gambar 6. Nominal to Polynominal

Tabel 6. Parameter pada Operator Nominal to Polynominal

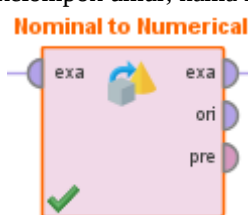
No	Parameter	Isi
1.	Attribute Filter Type	Subset
2.	Attributes	Tahun
3.	Coding Type	Unique Integer

Dari hasil pembacaan operator *Nominal to Polynominal* didapat informasi sebagai berikut.

Tabel 7. Statik Dataset

No	Uraian	Keterangan
1.	Record	971
2.	Special Attributes	1
3.	Regular Attributes	5
4.	Attributes :	
	Id	Integer, missing 0
	Nama Kabupaten Kota	Polynominal, missing 0
	Kelompok Umur	Polynominal, missing 0
	Jenis Kelamin	Polynominal, missing 0
	Jumlah Kasus	Polynominal, missing 0
	Tahun	Integer, missing 0

Nominal to Numerical, digunakan untuk mengubah tipe atribut yang bersifat non-numerik menjadi tipe numerik. Dalam dataset Kasus HIV/AIDS atribut jenis kelamin, kelompok umur, nama kabupaten kota, tahun dirubah menjadi numerik.



Gambar 7. Nominal to Numerical

Tabel 8. Parameter pada Pperator Nominal to Numerical

No	Parameter	Isi
1.	Attribute Filter Type	Subset
2.	Attributes	Jenis kelamin, kelompok umur, nama kabupaten kota, tahun
3.	Coding Type	Unique Integer

Dari hasil pembacaan operator *Nominal to Numerical* didapat informasi sebagai berikut.

Tabel 9. Statik Dataset

No	Uraian	Keterangan
1.	Record	971
2.	Special Attributes	1
3.	Regular Attributes	5
4.	Attributes :	
	Id	Integer, missing 0
	Nama Kabupaten Kota	Numeric, missing 0
	Kelompok Umur	Numeric, missing 0
	Jenis Kelamin	Numeric, missing 0
	Jumlah Kasus	Numeric, missing 0
	Tahun	Integer, missing 0

Data Mining

Dalam tahap ini algoritma yang digunakan adalah *k-means clustering*.



Gambar 8. K-Means Clustering

Parameter pada operator *K-Means Clustering* yang digunakan tampak pada tabel dibawah ini. Parameter pada operator *K-Means Clustering* ini menggunakan parameter default.

Tabel 10. Parameter pada Pperator K-Means Clustering

No	Parameter	Isi
1.	K	2-20

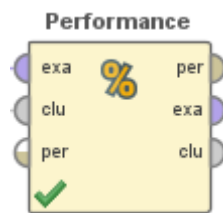
Dari hasil pembacaan operator *K-Means Clustering* dengan parameter *default* didapat informasi sebagai berikut.

Tabel 11. Hasil K-Means Clustering

iteration	Clusters	Measure_type	Main_criterion	Davies Bouldin
1	2	EuclideanDistance	Davies Bouldin	0.414
2	3	EuclideanDistance	Davies Bouldin	0.570
3	4	EuclideanDistance	Davies Bouldin	0.511
4	5	EuclideanDistance	Davies Bouldin	0.615
5	6	EuclideanDistance	Davies Bouldin	0.614
6	7	EuclideanDistance	Davies Bouldin	0.703
7	8	EuclideanDistance	Davies Bouldin	0.742
8	9	EuclideanDistance	Davies Bouldin	0.804
9	10	EuclideanDistance	Davies Bouldin	0.813

10	11	EuclideanDistance	Davies Bouldin	0.798
11	12	EuclideanDistance	Davies Bouldin	0.806
12	13	EuclideanDistance	Davies Bouldin	0.842
13	14	EuclideanDistance	Davies Bouldin	0.831
14	15	EuclideanDistance	Davies Bouldin	0.844
15	16	EuclideanDistance	Davies Bouldin	0.844
16	17	EuclideanDistance	Davies Bouldin	0.912
17	18	EuclideanDistance	Davies Bouldin	0.927
18	19	EuclideanDistance	Davies Bouldin	0.830
19	20	EuclideanDistance	Davies Bouldin	0.854

Selanjutnya digunakan operator *Performance* dengan metode *Davies-Bouldin Index (DBI)*. Untuk mengetahui nilai *DBI* nya.



Gambar 8. Cluster Distance Performance

Dari hasil pembacaan operator *Performance* didapat informasi sebagai berikut.

Tabel 12. Hasil Pembacaan Operator Performance

No	K	DBI
1.	2	0.414
2.	3	0.570
3.	4	0.511
4.	5	0.615
5.	6	0.614
6.	7	0.703
7.	8	0.742
8.	9	0.804
9.	10	0.813
10.	11	0.798
11.	12	0.806
12.	13	0.842
13.	14	0.831
14.	15	0.844
15.	16	0.844
16.	17	0.912
17.	18	0.927
18.	19	0.830
19.	20	0.854

Interpretation/Evaluation

Setelah melakukan perbandingan DBI dengan metode K-Means dari k-2 sampai k-20 yang terdapat pada tabel diatas, dapat dilihat bahwa cluster yang mendekati 0 yaitu k-2, fengan nilai DBI sebesar 0,414. Karena nilai k-2 merupakan nilai terkecil dibandingkan k lainnya, maka dapat disimpulkan bahwa k-2 dengan nilai 0,414 yang paling mendekati 0 merupakan hasil cluster terbaik.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil pengujian data menggunakan tools RapidMiner pada klasterisasi penyakit HIV/AIDS di Jawa Barat dengan menggunakan algoritma K-Means Clustering Setelah menghasilakn nilai DBI dengan metode K-Means dari k-2 sampai k-20 yang terdapat pada tabel diatas, dapat dilihat bahwa cluster yang mendekati 0 yaitu k-2, fengan nilai DBI sebesar 0,414. Karena nilai k-2 merupakan nilai terkecil dibandingkan k lainnya, maka dapat disimpulkan bahwa k-2 dengan nilai 0,414 yang paling mendekati 0 merupakan hasil *cluster* terbaik.

Saran

Beberapa saran yang harus diperhatikan dosen atau peneliti lain yang ingin mengimplementasikan lintasan belajar hasil penelitian ini berdasarkan pengalaman peneliti selama penelitian yaitu peneliti atau dosen harus memiliki keterampilan meneliti dan menulis artikel ilmiah yang baik, yang dapat dilihat dari pengalaman penelitian dan publikasi artikel ilmiah pada jurnal nasional maupun internasional.

Diharapkan dari penelitian ini dapat membuka paradigma berfikir dalam mendalami kasus penyebaran HIV/AIDS yang terjadi selama ini, serta menjadi bahan kajian dan memberi sumbangan pemikiran dan informasi dalam upaya pencegahan penyebaran HIV/AIDS di Provinsi Jawa Barat

DAFTAR PUSTAKA

- [1] E. Krisdayanti dan J. I. Hutasoit, "PENGARUH COPING STRATEGIES TERHADAP KESEHATAN MENTAL DAN KUALITAS HIDUP PENDERITA HIV/AIDS POSITIF," 2019.
- [2] S. Salami *dkk.*, "Studi Kualitatif Strategi Koping Penderita HIV AIDS di Kota Bandung," *Faletehan Health Journal*, vol. 8, no. 1, hlm. 22–30, 2021, [Daring]. Tersedia pada: www.journal.lppm-stikesfa.ac.id/ojs/index.php/FHJ
- [3] R. Indra, L. Sinaga, W. Saputra, H. Qurniawan, dan S. Tunas Bangsa, "Pengelompokan Jumlah Kasus Penyakit Aids Berdasarkan Provinsi Menggunakan Metode K-Means," 2021. [Daring]. Tersedia pada: <https://pusdatin.kemkes.go.id>
- [4] G. Gustientiedina, M. H. Adiya, dan Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 5, no. 1, hlm. 17–24, Apr 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [5] H. Priyatman, F. Sajid, dan D. Haldivany, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa," 2019.
- [6] Z. Nabila, A. Rahman Isnain, dan Z. Abidin, "ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS," *Jurnal Teknologi dan Sistem Informasi (JTSI)*, vol. 2, no. 2, hlm. 100, 2021, [Daring]. Tersedia pada: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [7] J. Nasir, "PENERAPAN DATA MINING CLUSTERING DALAM MENGELOMPOKAN BUKU DENGAN METODE K-MEANS," *Jurnal SIMETRIS*, vol. 11, no. 2, 2020.
- [8] M. Iqbal, "KLAUSTERISASI DATA JAMAAH UMROH PADA AULIYA TOUR & TRAVEL MENGGUNAKAN METODE K-MEANS CLUSTERING," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 5, no. 2, hlm. 97–104, Jun 2019, doi: 10.33330/jurteksi.v5i2.352.
- [9] K. Farouq dan H. Susilo, "KLAUSTERISASI VIRUS COVID-19 DI WILAYAH KABUPATEN LAMONGAN DENGAN METODE K-MEANS CLUSTERING."