JURNAL UPI

EDUTECH

Jurnal Teknologi Pendidikan

Journal homepage https://ejournal.upi.edu/index.php/edutech

EduTech JURNALTEKNOLOGI PENDIDIKAN

Penerapan Data Mining dalam Analisis Klasterisasi Penyebaran Penyakit HIV Menggunakan Algoritma K-Means

Bekti Dinar Cahyani, Irma Amelia Putri, Prawido Utomo, Detin Sofia
Institut Teknologi dan Bisnis Bina Sarana Global, Indonesia
Email: 1221130031@global.ac.id, 1221130106@global.ac.id, prawidoutomo@global.ac.id, detinsofia@global.ac.id

ABSTRACT

The Banten Health Office recorded cases of Human Immunodeficiency Virus (HIV) transmission reaching 2,100 people during the period from January to October 2024. The locations infected with HIV are spread across all districts/cities in Banten Province. However, the largest number is in the Greater Tangerang area which includes Tangerang City, Tangerang Regency and South Tangerang. The spread of HIV occurred in one of the Health Centers located in the Pasar Kemis area with an increasing number of cases in recent years. This study aims to carry out clustering by identifying the distribution of HIV cases in sub-districts with the highest and lowest levels of spread. This study applies data mining techniques, namely the Cross-Industry Standard Process for Data Mining (CRISP-DM) method with the K-Means Clustering algorithm that integrates the Elbow and Silhouette Coefficient methods by minimizing the SSE value. CRISP-DM includes 6 stages consisting of Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, Deployment. The analysis results obtained from each clustering with scatter plot visualization produced 4 clusters divided into 2 clusters with the highest and lowest cluster sub-districts. The highest sub-district is included in cluster_3 which is occupied by Sukamantri Sub-district with 11 HIV cases. While the lowest sub-district is included in cluster_0 which is occupied by Pasar Kemis Sub-district with 6 HIV cases.

ABSTRAK

Dinas Kesehatan Banten mencatat adanya kasus penularan Human Immunodeficiency Virus (HIV) yang mencapai 2.100 orang selama periode Januari sampai dengan Oktober 2024.

ARTICLE INFO

Article History:
Submitted/Received 5 Mei 2025
First Revised 12 Mei 2025
Accepted 25 Mei 2025
First Available online 01 Juni 2025
Publication Date 01 Juni 2025

Keyword:
HIV, CRISP-DM, K-Means, Data Mining

Adapun lokasi yang terinfeksi HIV tersebar di seluruh kabupaten/kota di Provinsi Banten. Namun jumlah terbanyak berada di daerah Tangerang Raya yang meliputi Kota Tangerang, Kabupaten Tangerang dan Tangerang Selatan. Penyebaran HIV terjadi pada salah satu Puskesmas yang terletak di daerah Pasar Kemis dengan jumlah kasus yang meningkat dalam beberapa tahun terakhir. Penelitian ini bertujuan untuk melakukan klasterisasi dengan mengidentifikasi persebaran kasus HIV pada kelurahan dengan tingkat penyebaran tertinggi dan terendah. Penelitian ini menerapkan teknik data mining yaitu metode Cross-Industry Standard Process for Data Mining (CRISP-DM) dengan algoritma K-Means Clustering yang mengintegrasikan metode Elbow dan Silhouette Coefficient dengan meminimalkan nilai SSE. CRISP-DM meliputi 6 tahap yang terdiri dari Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, Deployment. Hasil analisa yang diperoleh dari setiap klasterisasi dengan visualisasi scatter plot menghasilkan 4 cluster yang terbagi ke dalam 2 cluster dengan kelurahan cluster tertinggi dan terendah. Kelurahan tertinggi termasuk dalam cluster_3 yang ditempati oleh Kelurahan Sukamantri dengan 11 kasus HIV. Sementara Kelurahan terendah termasuk dalam cluster_0 yang ditempati oleh Kelurahan Pasar Kemis dengan 6 kasus HIV.

DOI: https://doi.org/10.17509/e.v24i2.84914
P-ISSN 0852-1190 E-ISSN 2502-0781

1. PENDAHULUAN

Kasus HIV/AIDS di Indonesia menunjukkan tren peningkatan. Berdasarkan data dari Sistem Informasi HIV/AIDS (SIHA) selama periode Januari hingga Desember 2023 tercatat sebanyak 57.299 orang dari 6.142.136 orang yang menjalani tes HIV dan sebanyak 46.370 orang mendapatkan pengobatan ARV (*Antiretroviral)*. Provinsi dengan jumlah kasus tertinggi adalah Jawa Timur, Jawa Barat, DKI Jakarta, Jawa Tengah, Sumatera Utara dan Bali (Kemenkes RI, 2022). Perkembangan kasus HIV di Provinsi Banten mengalami peningkatan yang semakin pesat dengan jumlah penduduk mencapai 12,47 juta jiwa. Dinas Kesehatan Banten mencatat 2.100 kasus baru HIV pada periode Januari hingga Oktober 2024. Dari jumlah tersebut, 9% atau 189 kasus dialami oleh Ibu Rumah Tangga (IRT). Sebaran kasus HIV pada Ibu Rumah Tangga (IRT) di Provinsi Banten tersebar merata di seluruh kabupaten/kota dengan jumlah kasus tertinggi di wilayah Tangerang Raya (Zuliansyah, 2024).

Penyebaran HIV terjadi pada Puskesmas yang terletak di daerah Pasar Kemis dengan jumlah kasus yang meningkat dalam beberapa tahun terakhir. Data menunjukkan adanya kasus HIV pada tahun 2015, 2019, dan 2020 masing-masing tercatat 1 kasus, pada 2021 tercatat 16 kasus, pada 2023 sebanyak 34 kasus, dan terus berlanjut hingga tahun 2025. Salah satu permasalahan yang terjadi pada pasien penularan HIV berasal dari pasangan dengan perilaku seks berisiko tinggi seperti sering berganti pasangan atau memiliki pasangan yang merupakan pelanggan pekerja seks komersial (PSK).

Pemerintah Kota Tangerang melalui Dinas Kesehatan terus berupaya mengoptimalkan layanan Perawatan, Dukungan dan Pengobatan (PDP) bagi Orang Dengan HIV/AIDS (ODHA). Layanan PDP untuk ODHA tersedia di Puskesmas Pasar Kemis yang menyediakan layanan Konseling, Tes *Viral Load* (VL), Tes CD4 dan Terapi *Antiretroviral* (ART). Untuk kasus yang memerlukan penanganan lebih intensif, pasien dapat mengajukan pemindahan data untuk dirujuk ke rumah sakit (Anggraeni, 2024b).

Pemerintah Provinsi Banten berupaya menanggulangi kasus HIV melalui pengobatan dan pencegahan dengan meningkatkan akses layanan HIV kepada masyarakat. Upaya ini mencakup *skrining* HIV yang salah satunya ditujukan kepada calon pengantin dan ibu hamil untuk mencegah penularan HIV *vertikal* dari ibu ke anak. Selain itu, *skrining triple* eliminasi untuk mendeteksi infeksi HIV, *sifilis, hepatitis* B, dan *skrining tuberkulosis* disediakan bagi ibu hamil. Layanan ini tersedia di seluruh puskesmas dan posyandu di Banten, bekerja sama dengan dokter atau bidan praktik mandiri di klinik wilayahnya serta melalui kolaborasi dengan kader untuk penjangkauan dan promosi kesehatan terkait pencegahan dan penularan HIV. Selain itu, obat pencegahan penularan HIV bernama PrEP (Profilaksis Pra Pajanan) disediakan bagi pasangan dengan HIV (ODHIV) atau individu dengan perilaku seks berisiko tinggi yang belum terinfeksi HIV (Anggraeni, 2024a).

Berdasarkan permasalahan yang telah diuraikan, penelitian ini bertujuan untuk melakukan klasterisasi guna mengidentifikasi persebaran kasus HIV pada kelurahan dengan kategori tingkat penyebaran tinggi maupun rendah namun memiliki karakteristik penyebaran yang serupa. Tujuannya adalah untuk mendukung layanan kesehatan dalam menyediakan informasi yang akurat sebagai dasar pengambilan kebijakan serta monitoring penyuluhan secara tepat sasaran dalam upaya pencegahan dan penanganan yang cepat dan efektif.

Penelitian ini menggunakan model *K-Means Clustering* yang dipadukan dengan metode *Elbow* dan *Silhouette Coefficient*. *K-Means* merupakan metode partisi klaster yang membagi data ke dalam sejumlah kelompok yang saling terpisah. Model ini dikenal karena kesederhanaannya serta kemampuannya dalam mengelompokkan data dalam skala besar secara efisien (Prabiantissa & Yuliastuti, 2021). Selain itu, *K-Means* mudah

DOI: https://doi.org/10.17509/e.v24i2.84914
p- ISSN 2528-1410 e- ISSN 2527-8045

diimplementasikan, memiliki kinerja yang cepat, fleksibel terhadap penyesuaian, dan telah banyak digunakan dalam berbagai aplikasi (Nur Amalia *et al*., 2024). Mekanisme kerja model ini adalah dengan meminimalkan *Sum of Squared Error* (SSE) antara setiap data dengan pusat klasternya (*centroid*) yang berjumlah k (Orisa & Ardita, 2020).

Berdasarkan penelitian yang dilakukan oleh (Kodratul Munawar & Irma Purnamasari, 2023) menggunakan atribut Nama kabupaten kota, kelompok umur, jenis kelamin, jumlah kasus dan tahun. Hasil penelitian menunjukkan pengelompokkan tingkat penyebaran HIV berdasarkan kelompok umur menggunakan algoritma K-Means Clustering diperoleh 2 cluster. Cluster_0 merupakan kategori rendah, sedangkan cluster_1 merupakan kategori tinggi yang terdapat pada kelompok umur 0-4 tahun di Kabupaten Subang dan Kota Cirebon, kelompok umur 20-24 tahun di Kabupaten Karawang dan Kota Cimahi, dan kelompok umur 25-49 tahun di Kabupaten Bogor, Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Cirebon, Kabupaten Indramayu, Kabupaten Subang, Kabupaten Purwakarta, Kabupaten Karawang, Kabupaten Bekasi, Kota Bogor, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, dan Kota Cimahi.

Penelitian dengan menggunakan algoritma yang sama dilakukan oleh (Wala, J., Herman, & Umar, 2024) menggunakan atribut Id, Age, Sex, Cp, Trestbps, Chol, Fbs, Restecg, Thalach, Exang, Oldpeak dan Slope. Hasil penelitian menunjukkan 5 cluster dengan cluster k1 berjumlah 27 pasien, k2 berjumlah 135 pasien, k3 berjumlah 15 pasien, dan k4 berjumlah 126 pasien. Analisis data menunjukan, cluster 1 (k1), cenderung terdiri dari pasien yang lebih tua, mayoritas laki-laki, menunjukan risiko tinggi dengan gejala nyeri dada parah, tekanan darah, dan kadar kolesterol tinggi. Sementara itu, cluster k2, k3, dan k4 menunjukkan risiko lebih rendah, dengan variasi respons terhadap aktivitas fisik.

Penelitian lainnya yang dilakukan oleh (Putra *et al*., 2022) menggunakan atribut Nama, Usia, Kelurahan Domisili, Hasil Pemeriksaan Awal dan Hasil Pemeriksaan Akhir. Hasil penelitian menunjukkan 3 klaster yang dikelompokkan berdasarkan kategori rendah, sedang, dan tinggi dengan nilai DBI sebesar -0,332. Pada klaster 0 dengan kategori rendah terdapat 3 kecamatan, klaster 1 dengan kategori sedang terdapat 1 kecamatan, klaster 2 dengan kategori tinggi terdapat 6 kecamatan. Hasil pengujian tersebut dapat diketahui bahwa usia rentan terhadap COVID-19 adalah 26 sampai dengan 45 tahun.

Secara umum, penelitian ini memiliki perbedaan dibandingkan dengan penelitian terdahulu yang terletak pada fokus penerapannya menggunakan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Metode ini menyediakan kerangka kerja yang terstruktur, fleksibel, dan komprehensif yang dapat disesuaikan dengan berbagai skenario data mining tanpa terikat pada teknologi tertentu.

1) Data Mining

Proses eksplorasi data dalam sebuah basis data yang besar guna menemukan pola atau tren tertentu yang dapat dimanfaatkan dalam pengambilan keputusan. Teknik ini telah digunakan sejak tahun 1990 dengan memanfaatkan data yang tersimpan dalam basis data (Rachman, F. H., Naim, I. J., & Aminah, 2023).

2) Clustering

Proses pengelompokan data seperti catatan, observasi, atau objek-objek yang memiliki karakteristik serupa ke dalam satu kelas. Sebuah *cluster* terdiri dari kumpulan data yang memiliki tingkat kemiripan tinggi antar anggotanya namun berbeda secara signifikan dengan data dalam *cluster* lainnya. Tujuan dari *clustering* adalah membagi keseluruhan data menjadi beberapa kelompok yang homogen, di mana kesamaan antar

DOI: https://doi.org/10.17509/e.v24i2.84914
P-ISSN 0852-1190 E-ISSN 2502-0781

data dalam satu kelompok dimaksimalkan, dan kesamaan dengan data dari kelompok lain diminimalkan (Swasika, R., Mukodimah, S., Susanto, F., Muslihudin, M., & Ipnuwati, 2023).

3) Algoritma K-Means

Model pembelajaran mesin tanpa pengawasan (*unsupervised*) yang sederhana dan banyak digunakan. *K-Means* merupakan salah satu metode analisis *cluster non-hierarki* yang bertujuan membagi data ke dalam satu atau lebih kelompok (*cluster*). Model ini populer karena kemudahannya serta kemampuannya dalam mengelompokkan data dalam jumlah besar maupun data yang mengandung *outlier* dengan cepat. Dalam proses *K-Means*, setiap data akan dimasukkan ke dalam salah satu *cluster* dan pada tiap iterasi, data tersebut dapat berpindah ke *cluster* lain hingga tercapai hasil akhir yang optimal (Rangkuti *et al*., 2023). Berikut cara kerja dari metode *K-Means* (Prabiantissa & Yuliastuti, 2021):

1) Tentukan jumlah *cluster* (kelompok) yang akan dibentuk.
2) Tentukan pusat *cluster* (*centroid*) awal, dapat dilakukan secara acak atau random.
3) Melakukan penghitungan terhadap jarak dari setiap objek ke *centroid* masing-masing *cluster* dengan rumus seperti berikut ini:

$$\boldsymbol{d(x,y) = \|x - y\| = \sqrt{\sum_{i-1}^{n}(x_i - y_i)^2}} \quad \textbf{(1)}$$

Keterangan:

d(x,y) = jarak antara x dan y

i = setiap data

n = jumlah data

x = data pusat klaster

y = data pusat atribut

x_i = data pada pusat klaster ke i

y_i = data pada setiap data ke i

4) Lakukan iterasi dengan menentukan *centroid* baru.
5) Ulangi iterasi sampai *centroid* tidak berubah.

4) Metode Elbow

Cara yang paling populer untuk menemukan jumlah *cluster* yang optimal dalam proses *clustering* agar hasil pengelompokan menjadi efektif (Setiawan, 2021). Metode Elbow menentukan kohesi dan pemisahan klaster yaitu keterkaitan data dalam klaster dan pemisahan antar klaster melalui perhitungan *Sum of Squared Error* (Vania & Nurina Sari, 2023). Dalam menghitung SSE dapat menggunakan rumus sebagai berikut:

$$\boldsymbol{SSE = \sum_{k=1}^{k}\sum_{x_i}|x_i - C_k|^2} \quad \textbf{(2)}$$

Keterangan:

K = Klaster ke-C

$\boldsymbol{x_i}$ = Jarak data pada objek ke-i

$\boldsymbol{C_k}$ = Pusat klaster ke-i

5) Metode Silhouette Coefficient

Metrik yang digunakan untuk menilai seberapa mirip suatu data dengan anggota lain dalam klasternya dan dihitung secara individual untuk setiap data. Nilai koefisien yang mendekati 1 menunjukkan bahwa pengelompokan data dalam klaster tersebut memiliki kualitas yang baik. Sebaliknya, nilai yang mendekati -1 mengindikasikan bahwa pengelompokan tersebut kurang tepat atau buruk. Semakin besar koefisien Silhouette, semakin baik klaster tersebut (Rahmawati, T., Wilandari, Y., & Kartikasari, 2024). Menurut Rousseeuw (1987), kriteria pengukuran *Silhouette Coefficient* sebagai berikut:

DOI: https://doi.org/10.17509/e.v24i2.84914
p- ISSN 2528-1410 e- ISSN 2527-8045

Tabel 1. Ukuran *Silhouette Coefficient*

Silhouette Coefficient	**Interpretasi yang disesuaikan**
0,7 < SC ≤ 1,0	Susunan sangat baik
0,5 < SC ≤ 0,7	Susunan baik
0,25 < SC ≤ 0,5	Susunan lemah
SC ≤ 0,25	Susunan buruk

$$\boldsymbol{SC = maks_k SI\ (k)} \quad \textbf{(3)}$$

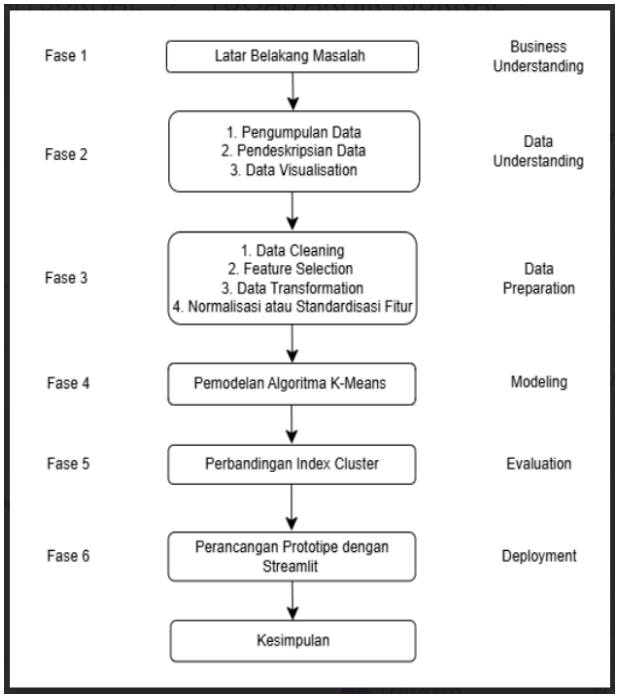
Keterangan:
SC = *Silhouette Coefficient*
SI = *Silhouette Index Global*
k = Jumlah klaster

6) SSE (Sum of Square Error)

Nilai SSE yang dievaluasi untuk menentukan apakah nilai minimum telah tercapai. Jika nilai SSE belum mencapai nilai minimum (mendekati 0), maka proses akan diulang kembali ke langkah penentuan nilai K (atau K+1) dan seterusnya, hingga ditemukan nilai SSE yang minimum (mendekati 0). Nilai selisih SSE terbesar (penurunan signifikan) menunjukkan nilai K *cluster* yang paling optimal dengan akurasi terbaik (Refialy *et al*., 2021). Karena semakin kecil nilai SSE, semakin seragam data dalam setiap *cluster*, dan semakin baik hasil *clustering* nya (Nur Amalia *et al*., 2024).

2. METODE

Penelitian ini menerapkan teknik *data minin*g menggunakan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM) meliputi 6 *fase* yang terdiri dari (Rohmah Zaidah *et al*., 2021):



Gambar 1. Alur CRISP-DM

DOI: https://doi.org/10.17509/e.v24i2.84914
P-ISSN 0852-1190 E-ISSN 2502-0781

1. Business Understanding
   Menentukan tujuan dan kebutuhan penelitian secara keseluruhan dengan mengidentifikasi permasalahan (latar belakang masalah) dan menyiapkan strategi awal untuk mencapai tujuan.
2. Data Understanding
   Proses pengumpulan data yang akan dianalisis, pendeskripsian data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal serta melakukan visualisasi data dalam bentuk diagram.
3. Data Preparation
   Mempersiapkan dan mengumpulkan data yang akan digunakan untuk keseluruhan *fase* berikutnya. Proses yang dilakukan pada *fase* ini meliputi data *cleaning* (pembersihan data) dengan mengecek duplikasi data, menghapus duplikasi dan mengecek *missing value* selanjutnya pemilihan fitur terhadap data dan transformasi data dengan mengubah data teks menjadi numerik serta melakukan normalisasi atau standarisasi fitur sehingga data siap untuk dimodelkan.
4. Modeling
   Implementasi algoritma *K-Means Clustering* menggunakan data yang telah disiapkan untuk kebutuhan analisis dengan bantuan *tools Google Colaboratory*. *Fase* ini menerapkan metode *Elbow* berdasarkan perhitungan nilai *Sum of Square Error*.
5. Evaluation
   Evaluasi pada algoritma yang digunakan dengan menerapkan metode *silhouette coefficient* kemudian melakukan perbandingan *index cluster* berdasarkan SSE dan *Betweenss* untuk mendapatkan kualitas dan efektivitas sebelum memasuki *fase* penyebaran.
6. Deployment
   Algoritma yang dihasilkan digunakan untuk perancangan *prototype* menggunakan *streamlit* dengan membuat situs *website* sederhana yang dapat diakses melalui internet.

3. HASIL DAN PEMBAHASAN

Berdasarkan metode CRISP-DM, penelitian ini terdiri atas beberapa tahapan pada hasil dan pembahasan yaitu sebagai berikut:

3.1. Business Understanding (Pemahaman Bisnis)

Human Immunodeficiency Virus (HIV) adalah *virus* yang melemahkan sistem kekebalan tubuh dengan menyerang sel darah putih, sehingga penderitanya rentan terhadap penyakit seperti TBC, *infeksi*, dan kanker. *Virus* ini menular melalui cairan tubuh seperti darah, ASI, air mani, dan cairan vagina, serta dari ibu ke anak. HIV tidak menular melalui kontak biasa seperti ciuman atau berbagi makanan. HIV dapat dicegah dan dikendalikan dengan terapi *antiretroviral* (ART). Jika tidak diobati, HIV dapat berkembang menjadi AIDS. WHO mendefinisikan HIV stadium lanjut (AHD) sebagai jumlah sel CD4 di bawah 200 sel/mm^3 atau masuk stadium 3/4 pada orang dewasa dan remaja. Anak di bawah 5 tahun dengan HIV juga digolongkan dalam AHD (WHO, 2024).

3.2. Data Understanding (Pemahaman Data)

Penelitian ini menggunakan *dataset* yang bersumber dari Puskesmas Pasar Kemis. *Dataset* HIV memiliki data yang berjumlah 88 baris dan 6 kolom (atribut) di dalamnya. Atribut yang terdapat dalam data HIV meliputi Umur, Jenis Kelamin, Kecamatan, Kelurahan, Tanggal Register, dan Status ODHIV. Data tersebut menunjukkan adanya kasus HIV dari 2015 hingga tahun 2025.

DOI: https://doi.org/10.17509/e.v24i2.84914
p- ISSN 2528-1410 e- ISSN 2527-8045

Tabel 2. *Dataset* HIV

Umur	Jenis Kelamin	Kecamatan	Kelurahan	Tahun Register	Status ODHIV
45	Laki-laki	Pasar Kemis	Sukamantri	2023	Meninggal
17	Perempuan	Pasar Kemis	Pangadegan	2020	Gagal follow up
37	Laki-laki	Pasar Kemis	Pangadegan	2021	ODHIV sedang pengobatan
28	Perempuan	Pasar Kemis	Pangadegan	2021	ODHIV sedang pengobatan
39	Laki-laki	Rajeg	Sukamanah	2021	ODHIV sedang pengobatan
30	Laki-laki	Pasar Kemis	Sukaasih	2021	ODHIV sedang pengobatan
50	Laki-laki	Pasar Kemis	Sukamantri	2021	ODHIV sedang pengobatan
29	Laki-laki	Pasar Kemis	Sukamantri	2021	Meninggal
27	Laki-laki	Pasar Kemis	Pangadegan	2021	Meninggal
29	Laki-laki	Periuk	Periuk	2021	ODHIV sedang pengobatan
28	Laki-laki	Pasar Kemis	Pasar Kemis	2021	ODHIV sedang pengobatan
52	Laki-laki	Pasar Kemis	Kutabumi	2022	Meninggal
41	Perempuan	Pasar Kemis	Sukaasih	2022	ODHIV sedang pengobatan
35	Laki-laki	Padas	Kedungprahu	2022	Gagal follow up
26	Laki-laki	Pasar Kemis	Pasar Kemis	2022	ODHIV sedang pengobatan
35	Laki-laki	Cikupa	Suka Damai	2022	Meninggal
18	Laki-laki	Cikupa	Pasir Gadung	2022	Meninggal
49	Laki-laki	Sindang Jaya	Wana Kerta	2022	Gagal follow up
31	Perempuan	Sindang Jaya	Wana Kerta	2022	Meninggal
42	Perempuan	Pasar Kemis	Sukaasih	2022	ODHIV sedang pengobatan

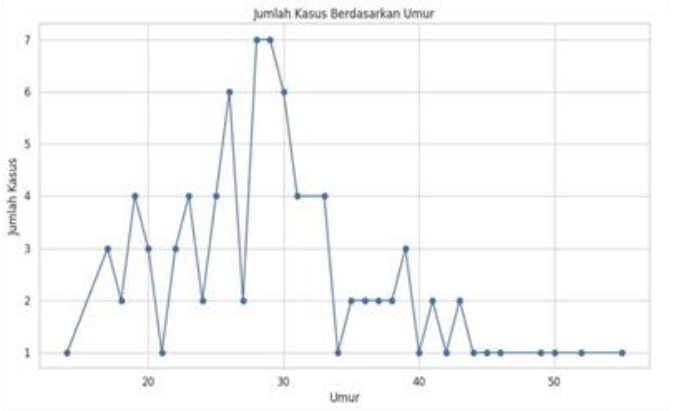
Penelitian ini menyajikan visualisasi data dalam bentuk diagram untuk mempermudah pemahaman dari *dataset* tersebut. Terdapat 4 jenis visualisasi yang digunakan yaitu:

1. Diagram *Wordcloud*

Gambar 2. Diagram *Wordcloud* Kelurahan

Berdasarkan gambar 2, menunjukkan jumlah penemuan kasus ODHIV pada setiap kelurahan. Ditemukan 3 kelurahan dengan jumlah kasus HIV terbesar dibandingkan kelurahan lainnya yaitu kelurahan Sukamantri, Sindang Sari dan Pangadegan.

2. Diagram Garis



Gambar 3. Diagram Garis Umur

Berdasarkan gambar 3, menunjukkan adanya peningkatan jumlah kasus HIV dari umur remaja hingga dewasa. Jumlah kasus menurun secara bertahap dan stabil pada angka rendah di umur 40 hingga 50. umur 19, 23, 25, 31 dan 33 merupakan kelompok umur remaja hingga dewasa dengan jumlah kasus tertinggi, hal tersebut menunjukkan bahwa kelompok umur ini mungkin lebih aktif secara sosial sehingga lebih mudah

DOI: https://doi.org/10.17509/e.v24i2.84914
P-ISSN 0852-1190 E-ISSN 2502-0781

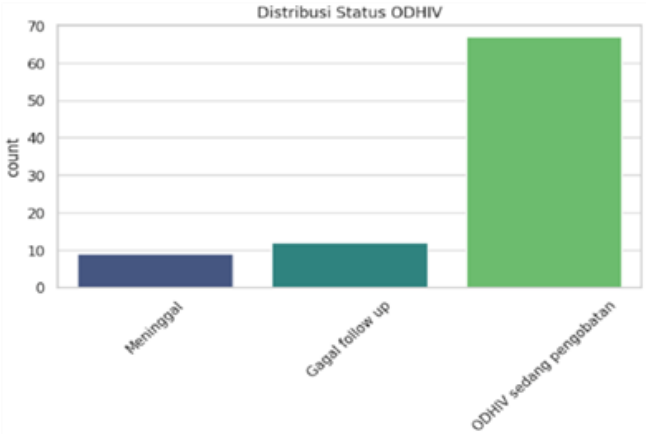
terpapar HIV. Berdasarkan penjelasan tersebut, dapat disimpulkan bahwa kelompok umur remaja hingga dewasa perlu mendapatkan monitoring penyuluhan secara tepat sasaran dalam upaya pencegahan dan penanganan yang cepat dan efektif berdasarkan umur.

3. Diagram *Pie*

Gambar 4. Diagram *Pie* Jenis Kelamin

Berdasarkan gambar 4, menunjukkan *persentase* jumlah kasus ODHIV yang ditemukan pada periode tahun 2015-2025, dimana kasus HIV lebih banyak terjadi pada jenis kelamin laki-laki sebesar 77.27% dan perempuan sebesar 22.73%.

4. Diagram Batang



Gambar 5. Diagram Batang Status ODHIV

Berdasarkan gambar 5, menunjukkan perkembangan jumlah Status Orang Dengan HIV (ODHIV) yang ditemukan pada periode tahun 2015-2025 terdapat peningkatan signifikan pada status ODHIV sedang pengobatan berkisar 67 orang. Akan tetapi status ODHIV meninggal terjadi sedikit penurunan berkisar 9 orang. Dan status ODHIV gagal follow up tidak ada kenaikan berkisar 12 orang.

3.3. Data Preparation (Persiapan Data)

Data dalam bentuk *csv* diolah menggunakan tools *Google Colaboratory.* Langkah awal dalam proses ini adalah memeriksa redundansi data, di mana ditemukan 2 data yang terduplikasi, kemudian dilakukan pembersihan sehingga tidak ada lagi data yang duplikat. Selanjutnya memeriksa *missing value* pada data HIV yang diolah dan data tersebut tidak memiliki nilai kosong dikarenakan sebelumnya penulis telah mengisi (*imputasi*) data secara manual pada 2 atribut yaitu Kecamatan dan Kelurahan.

Penulis melakukan pemilihan fitur terhadap atribut data yang dianggap tidak penting. Atribut data dianggap tidak relevan dengan penelitian karena tidak memengaruhi proses analisis data. Proses tersebut dilakukan dengan menghapus beberapa atribut secara

DOI: https://doi.org/10.17509/e.v24i2.84914
p- ISSN 2528-1410 e- ISSN 2527-8045

manual seperti NIK/No. Identitas, Nama Pasien, Tanggal Lahir, Alamat Provinsi, Alamat Kabupaten, Alamat Domisili, No Rekam Medik, Kunjungan Terakhir, dan No Reg Nas.

Proses selanjutnya, melakukan transformasi data secara manual pada atribut Tanggal Register menjadi Tahun Register. Dari 6 atribut data HIV akan dipilih 4 atribut untuk dilakukan transformasi *encoding* data *kategorikal* berbentuk objek (*string* atau *teks*) ke dalam nilai *numerik* menggunakan teknik *label encoding*.

	Umur	Jenis Kelamin	Kecamatan	Kelurahan	Tahun Register	Status ODHIV
0	45	0	16	33	2023	1
1	17	1	16	17	2020	0
2	37	0	16	17	2021	2
3	28	1	16	17	2021	2
4	39	0	19	32	2021	2

Gambar 6. Transformasi *Encoding*

Berdasarkan gambar 6, menampilkan pengelompokan hasil *konversi* nilai transformasi *encoding*. Dengan keterangan atribut masing-masing sebagai berikut: Jenis Kelamin 0 = Laki-laki dan 1 = Perempuan. Kecamatan 16 = Pasar Kemis dan 19 = Rajeg. Kelurahan 33 = Sukamantri, 17 = Pangadegan, 32 = Sukamanah. Status ODHIV 0 = Gagal follow up, 1 = Meninggal, 2 = ODHIV sedang pengobatan.

Proses selanjutnya, melakukan normalisasi atau *feature scaling* yaitu proses transformasi nilai setiap variabel data sehingga semua *feature* hasil transformasi terpadatkan dalam jangkauan rentang nilai yang seragam dan menunjukkan baris data dengan jangkauan angka yang berbeda-beda untuk setiap *feature*. Penjelasan tersebut dapat dilihat pada gambar 7.

Data setelah Standarisasi:

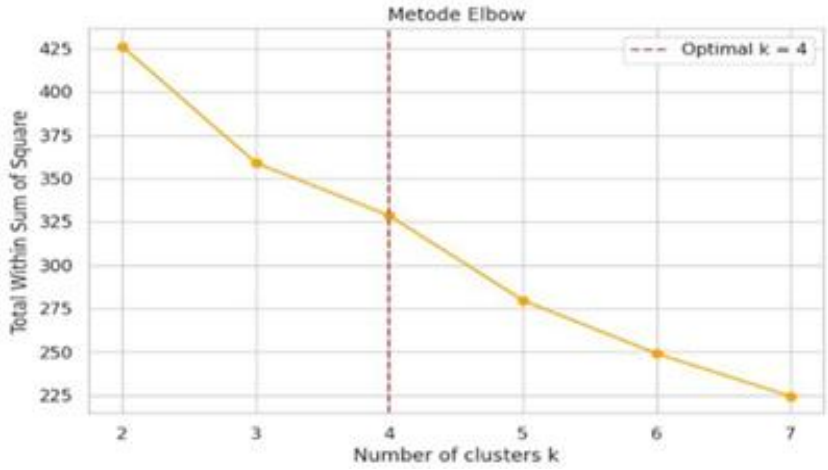
	Umur	Jenis Kelamin	Kecamatan	Kelurahan	Tahun Register	Status ODHIV
0	1.699551	-0.550482	0.195062	1.053656	0.272848	-0.858307
1	-1.515311	1.816590	0.195062	-0.477316	-1.797588	-2.251031
2	0.781019	-0.550482	0.195062	-0.477316	-1.107443	0.534417
3	-0.252329	1.816590	0.195062	-0.477316	-1.107443	0.534417
4	1.010652	-0.550482	0.724809	0.957970	-1.107443	0.534417

Gambar 7. Normalisasi atau Standardisasi Fitur

3.4. Modeling (Pemodelan)

Penulis akan menentukan jumlah *cluster* paling optimal dalam model *K-Means* menggunakan metode *Elbow* berdasarkan perhitungan nilai *Sum of Square Error* dengan menggunakan alat bantu bahasa *pemrograman python.* Metode ini berprinsip bahwa semakin besar nilai jumlah *cluster* K maka nilai SSE akan semakin kecil.

DOI: https://doi.org/10.17509/e.v24i2.84914
P-ISSN 0852-1190 E-ISSN 2502-0781



Gambar 8. Metode *Elbow*

Berdasarkan gambar 8, grafik ini menunjukkan jumlah *cluster* optimal adalah 4 *cluster* dengan titik dimana siku mulai membentuk garis lurus dan stabil berada pada C=4. Selanjutnya, dilakukan proses *clustering* dengan jumlah *cluster* yang terbentuk pada data berdasarkan perhitungan nilai *Sum of Square Error* yaitu jumlah jarak kuadrat antara tiap titik data dan *centroid cluster*.

```
K-means clustering with 4 clusters of sizes:
(0: 22, 1: 22, 2: 24, 3: 18)

Cluster means:
       Umur  Jenis Kelamin  Kecamatan  Kelurahan  Tahun Register  Status ODHIV
0 -0.701157      -0.550482  -0.134023   0.697009        0.398329      0.534417
1  0.133872      -0.335294   0.010453   0.201183       -0.636689     -1.301446
2  0.269129      -0.550482  -0.077169  -0.883981        0.071556      0.534417
3  0.334511       1.816590   0.253923   0.080051        0.196165      0.224923

Within cluster sum of squares by cluster:
(1: np.float64(45.51592348146849), 2: np.float64(146.66795110105403), 3: np.float64(73.29299595168224), 4: np.float64(63.055559190810245))

(between_SS / total_SS = 36.3 %)

Total within Sum of Squares (tot.withinss): 328.53
Between-Cluster Sum of Squares (betweenss): 187.47
```

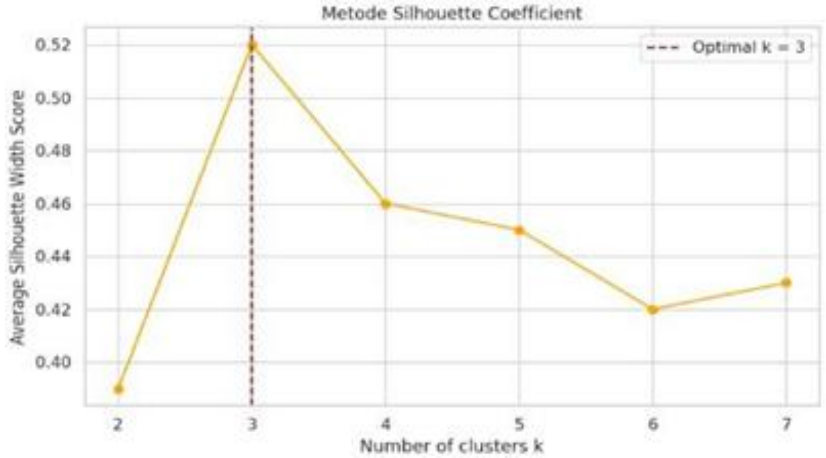
Gambar 9. Hasil *Clustering* C=4

Berdasarkan gambar 9, menunjukkan hasil *clustering* data yang telah di-skala dengan jumlah nilai C=4. Dimana jumlah anggota atau ukuran dari cluster_0 terdapat 22 anggota, cluster_1 terdapat 22 anggota, cluster_2 terdapat 24 anggota dan cluster_3 terdapat 18 anggota. Hasil *clustering* ini memperlihatkan bahwa nilai *betweenss* atau perhitungan jarak antara *cluster* adalah 187.47. Adapun total jumlah kuadrat di dalam model *clustering* yaitu total dari *withinss cluster* atau sering disebut nilai *Sum of Square Error* (SSE) adalah 328.53.

3.5. Evaluation (Pengujian)

Berdasarkan evaluasi menggunakan metode s*ilhouette coefficient, validasi* jumlah *cluster* dilakukan untuk menentukkan nilai k yang tepat dengan uji kelayakan dan kualitas klasterisasi model berdasarkan perhitungan nilai *Sum of Square Error*.

DOI: https://doi.org/10.17509/e.v24i2.84914
p- ISSN 2528-1410 e- ISSN 2527-8045



Gambar 10. Metode *Silhouette Coefficient*

Berdasarkan gambar 10, grafik ini menunjukkan ukuran kualitas jumlah *cluster* berada pada titik optimal terletak pada C=3 dimana nilai *koefisien silhouette* adalah 0.52 lebih besar daripada nilai jumlah *cluster* lainnya. Rentang nilai yang dihasilkan termasuk kriteria susunan baik. Selanjutnya, dilakukan proses *clustering* dengan jumlah *cluster* yang terbentuk pada data berdasarkan perhitungan nilai *Sum of Square Error* yaitu jumlah jarak kuadrat antara tiap titik data dan *centroid cluster.*

```
K-means clustering with 3 clusters of sizes:
{0: 49, 1: 19, 2: 18}

Cluster means:
       Umur  Jenis Kelamin  Kecamatan  Kelurahan  Tahun Register  Status OOHIV
0 -0.142199      -0.550482  -0.053595  -0.071140        0.117918      0.534417
1  0.152550      -0.301316  -0.158103   0.101834       -0.199357     -1.737922
2  0.226073       1.816590   0.312784   0.086167       -0.110566      0.379670

Within cluster sum of squares by cluster:
{1: np.float64(183.6939224777104), 2: np.float64(93.3019847431545), 3: np.float64(82.08954977907995)}

(between_SS / total_SS = 30.4 %)

Total Within Sum of Squares (tot.withinss): 359.09
Between-Cluster Sum of Squares (betweenss): 156.91
```

Gambar 11. Hasil *Clustering* C=3

Berdasarkan gambar 11, menunjukkan hasil *clustering* data yang telah di-skala dengan jumlah nilai C=3. Dimana jumlah anggota atau ukuran dari cluster_0 terdapat 49 anggota, cluster_1 terdapat 19 anggota, cluster_2 terdapat 18 anggota. Hasil *clustering* ini memperlihatkan bahwa nilai *betweenss* atau perhitungan jarak antara *cluster* adalah 156.91. Adapun total jumlah kuadrat di dalam model *clustering* yaitu total dari *withinss cluster* atau sering disebut nilai *Sum of Square Error* (SSE) sebesar 359.09.

Tabel 3. Perbandingan *Index Cluster*

Metode	Cluster	SSE/Tot.Withinss	Betweenss
Elbow	4	328.53	187.47
Silhouette	3	359.09	156.91

Berdasarkan tabel 3, menunjukkan perbedaan dalam penentuan jumlah *cluster* optimal. Grafik metode *Elbow* sejumlah *cluster* optimal pada titik C=4 dan grafik metode *Silhouette Coefficient* sejumlah *cluster* optimal pada titik C=3. Dapat disimpulkan bahwa metode *Elbow* pada titik C=4 lebih unggul daripada *Silhouette Coefficient* dalam menentukan jumlah *cluster* yang optimal. Karena pada jumlah *cluster Elbow* diperoleh nilai SSE/*Tot.Withinss* lebih rendah yaitu 328.53 dan nilai *Betweenss* lebih tinggi yaitu 187.47.

DOI: https://doi.org/10.17509/e.v24i2.84914
P-ISSN 0852-1190 E-ISSN 2502-0781

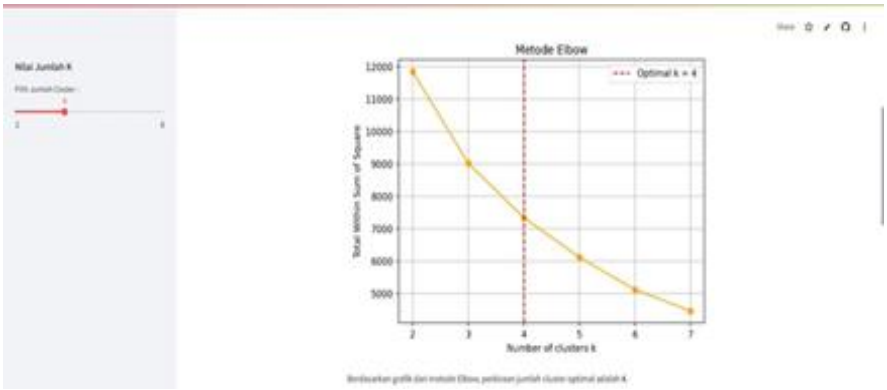
3.6. Deployment (Penyebaran)

Perancangan *prototype* (*interface*) dilakukan menggunakan *streamlit* dengan menampilkan model *K-Means Clustering* ke dalam file *streamlit* agar dapat diakses melalui *web browser* kemudian dilakukan *deployment* sehingga dapat diakses melalui *internet*. Situs web sederhana yang dibuat hanya menghasilkan distribusi dan karakteristik dari setiap *cluster* dengan memahami pola dan tren yang ada dalam data. Aplikasi *prototype* ini dapat membantu Dinas Kesehatan dalam memantau wilayah berisiko tinggi secara visual dan interaktif, sehingga dapat mempercepat proses pengambilan keputusan dalam intervensi kesehatan masyarakat.

File Python dengan ekstensi .py digunakan untuk membuat tampilan situs *web* sederhana pada *streamlit*. *Import streamlit* digunakan untuk memanggil *library,* sementara pandas digunakan untuk memuat kembali *dataset* ke dalam *streamlit*. Perintah st.header("Isi Dataset") dan st.write(HIV) berfungsi untuk menampilkan judul serta isi dataset di *streamlit*. Penjelasan tersebut dapat dilihat pada gambar 12.

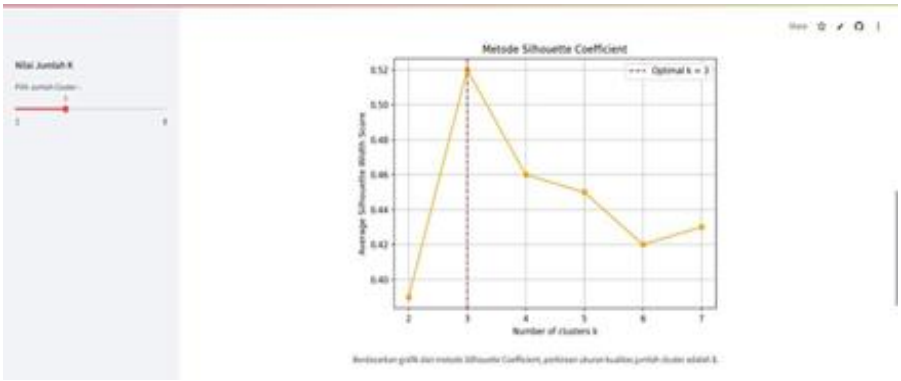
Gambar 12. *Dataset* dan Jumlah K pada *Streamlit*

Perintah *st.pyplot(fig)* berfungsi untuk memanggil dan menampilkan visualisasi grafik *elbow* dan *silhouette coefficient* dari model *K-Means Clustering* pada *streamlit*. Berdasarkan grafik *elbow* menghasilkan jumlah *cluster* optimal adalah 4 seperti pada gambar 13 dan grafik *silhouette coefficient* menghasilkan ukuran kualitas jumlah *cluster* adalah 3 dengan rentang nilai yang dihasilkan termasuk kriteria susunan baik dapat dilihat pada gambar 14.



Gambar 13. Metode *Elbow* pada *Streamlit*

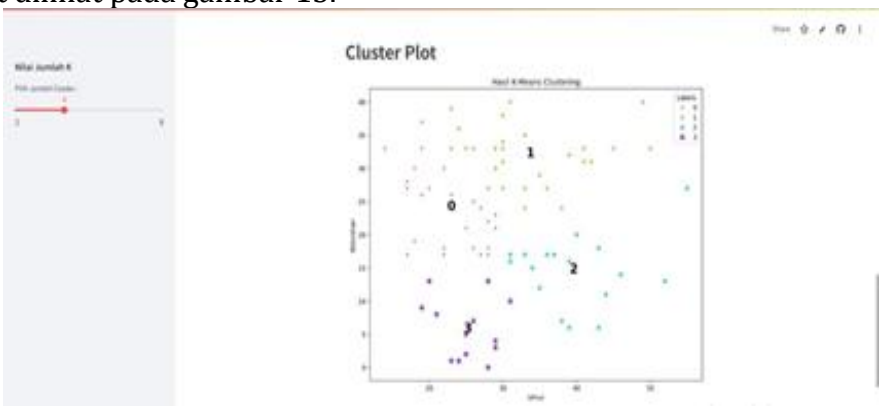
DOI: https://doi.org/10.17509/e.v24i2.84914
p- ISSN 2528-1410 e- ISSN 2527-8045



Gambar 14. Metode *Silhouette Coefficient* pada *Streamlit*

Setelah beberapa *cluster* terbentuk dari metode *elbow* dan *silhouette coefficient,* selanjutnya menentukan jumlah *cluster* dimana "Nilai Jumlak K" dapat di *custom* dalam bentuk *sliding* menggunakan *st.sidebar.slider* di dalam *interface streamlit*. Proses *clustering def k_means(n_clust)* dijalankan untuk memasukkan fungsi *cluster* berdasarkan nilai *slider*, dimana *n_clust* untuk memanggil fungsi dari model.

Hasil akhir *clustering* divisualisasikan menggunakan *scatter plot* dalam bentuk *figure* berdasarkan atribut umur dan kelurahan. *st.header('Cluster Plot')* untuk memanggil *label* di *streamlit.* Visualisasi *scatter plot* pada *streamlit* menampilkan distribusi titik data dalam *cluster* yang berbeda, dengan setiap *cluster* diwakili oleh warna yang berbeda. Penjelasan tersebut dapat dilihat pada gambar 15.



Gambar 15. *Cluster Plot* pada *Streamlit*

Berdasarkan gambar 15, menunjukkan terdapat 4 *cluster* yaitu cluster_3 dari Kelurahan Sukamantri dengan 11 kasus HIV tertinggi disertai Umur tertinggi penderita HIV di kelurahan ini termasuk kelompok umur dewasa yaitu 50 tahun. Selain itu, kasus serupa juga terjadi pada kelompok umur remaja yaitu 14 tahun, di kelurahan yang sama. Selanjutnya pada cluster_2 terdapat Kelurahan Sindang Sari dengan 9 kasus HIV tertinggi kedua setelah Kelurahan Sukamantri disertai Umur tertinggi penderita HIV di kelurahan ini termasuk kelompok umur dewasa yaitu 55 tahun. Kasus serupa juga terjadi pada kelompok umur remaja yaitu 17 tahun, di kelurahan ini. Selanjutnya pada cluster_1 terdapat Kelurahan Pangadegan dengan 9 kasus HIV tingkat menengah disertai Umur tertinggi penderita HIV di kelurahan ini termasuk kelompok umur dewasa yaitu 37 tahun. Kasus serupa juga terjadi pada kelompok umur remaja yaitu 17 tahun, di kelurahan ini. Kemudian cluster_0 terdapat Kelurahan Pasar Kemis dengan 6 kasus HIV terendah disertai Umur tertinggi penderita HIV di kelurahan ini termasuk kelompok umur dewasa, mulai dari 22 hingga 43 tahun.

DOI: https://doi.org/10.17509/e.v24i2.84914
P-ISSN 0852-1190 E-ISSN 2502-0781

4. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis *clustering* data yang telah diolah menggunakan *K-Means* dan penelitian yang dilakukan di Puskesmas Pasar Kemis, dapat disimpulkan sebagai berikut:

1. Analisis clustering berdasarkan kelurahan dan umur pada kasus HIV di Puskesmas Pasar Kemis mengidentifikasi 4 klaster dengan distribusi geografis dan demografis yang berbeda. Kelurahan Sukamantri (Cluster_3) tercatat memiliki jumlah kasus tertinggi, Kelurahan Sindang Sari (Cluster_2) dan Kelurahan Pangadegan (Cluster_1) tercatat memiliki jumlah kasus yang sama, sementara Kelurahan Pasar Kemis (Cluster_0) memiliki jumlah kasus terendah. Penyebaran kasus di setiap Kelurahan menunjukkan pola konsisten antara kelompok umur remaja hingga dewasa dengan umur tertinggi mencapai 55 tahun dan umur terendah 14 tahun. Selain itu, dari penelitian ini menunjukkan bahwa Kelurahan Sukamantri (Cluster_3) memiliki jumlah kasus HIV tertinggi dengan 11 kasus, sementara Kelurahan Pasar Kemis (Cluster_0) memiliki jumlah kasus terendah dengan 6 kasus.
2. Analisis proses *clustering* menggunakan metode *elbow* dan *silhouette coefficient* berdasarkan perhitungan nilai *Sum of Square Error* dengan penentuan nilai SSE/*Tot.withinss* dan *Betweenss*. Penelitian ini menunjukkan bahwa pada grafik metode *elbow,* titik *cluster* optimal berada pada C=4 dan lebih unggul daripada *silhouette coefficient k*arena pada jumlah *cluster elbow* diperoleh nilai SSE/*Tot.withinss* lebih rendah yaitu 328.53 dan nilai *Betweenss* lebih tinggi yaitu 187.47. Sementara, pada grafik metode *silhouette coefficient*, titik *cluster* optimal dari model yang terbentuk menghasilkan rentang nilai yang termasuk kriteria susunan baik dengan C=3, dimana nilai SSE/*Tot.withinss* sebesar 359.09 dan nilai *Betweenss* sebesar 156.91.

Berdasarkan hasil analisis *clustering* data yang telah diolah menggunakan *K-Means* dan penelitian yang dilakukan di Puskesmas Pasar Kemis, maka terdapat beberapa saran untuk penelitian selanjutnya sebagai berikut:

1. Memperluas area cakupan ke beberapa puskesmas sehingga data daerah sebaran yang dihasilkan lebih luas.
2. Menggunakan algoritma *data mining* klasterisasi lainnya untuk memperoleh perbandingan hasil yang lebih akurat serta lebih efektif dan efisien.
3. Melakukan teknik pra pengolahan data lainnya sehingga mendapatkan hasil data yang lebih baik.

5. PERNYATAAN PENULIS

Penulis menyatakan bahwa tidak terdapat konflik kepentingan terkait penerbitan artikel ini. Penulis menegaskan bahwa naskah artikel bebas dari plagiarisme.

6. REFERENSI

Anggraeni, D. (2024a). *Dinkes Berikan Layanan HIV Gratis di Seluruh Faskes Kota Tangerang*. SEKRETARIAT DAERAH KOTA TANGERANG. Diakses dari: https://setda.tangerangkota.go.id/berita/dinkes-berikan-layanan-hiv-gratis-di-seluruh-faskes-kota-tangerang

Anggraeni, D. (2024b). *Hari AIDS Sedunia, Berikut Layanan Kesehatan untuk ODHA di Kota Tangerang*. Pemerintah Kota Tangerang. Diakses dari: https://www.tangerangkota.go.id/berita/detail/47831/hari-aids-sedunia-berikut-layanan-kesehatan-untuk-odha-di-kota-tangerang

Kemenkes RI. (2022). *Perkembangan Hiv Aids Dan Penyakit Infeksi Menular Seksual (Pims) Triwulan IV Tahun 2022*. Kemenkes RI.

DOI: https://doi.org/10.17509/e.v24i2.84914
p- ISSN 2528-1410 e- ISSN 2527-8045

https://siha.kemkes.go.id/portal/files_upload/Laporan_TW_3_2022.pdf

Kodratul Munawar, K., & Irma Purnamasari, A. (2023). Implementasi Algoritma K-Means Clustering Pada Klasterisasi Kasus Hiv Di Jawa Barat. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(2), 1092–1099. https://doi.org/10.36040/jati.v7i2.6372

Nur Amalia, I., Umaidah, Y., & Mayasari, R. (2024). Penerapan Data Mining Untuk Klasterisasi Daerah Rawan Penyakit Menular Di Kabupaten Karawang Dengan Menggunakan Algoritma K-Means. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(4), 5582–5591. https://doi.org/10.36040/jati.v8i4.9953

Orisa, M., & Ardita, M. (2020). Web Usage Mining Menggunakan Algoritma Clastering K-Mean. *Jurnal Teknologi Informasi dan Terapan*, *8*(1), 60–64. https://doi.org/10.25047/jtit.v8i1.179

Prabiantissa, C. N., & Yuliastuti, G. E. (2021). Prediksi Pergerakan Ikan Di Pesisir Pulau Madura Menggunakan Metode Gaussian Mixture Model Dan K-Means Clustering. *Jurnal Teknologi Informasi dan Terapan*, *8*(2), 121–128. https://doi.org/10.25047/jtit.v8i2.244

Putra, A. Z., Pinem, R. W., Silalahi, S., Gulo, F., & Liukhoto, J. A. A. (2022). Classification of Covid-19 Patient Spread Rate By Age and Region With K-Means Algorithm. *SinkrOn*, *7*(3), 1085–1989. https://doi.org/10.33395/sinkron.v7i3.11603

Rachman, F. H., Naim, I. J., & Aminah, F. F. N. (2023). *Data Mining*. Literasi Nusantara. https://ipusnas2.perpusnas.go.id/book/b8ba1b73-665c-4a57-a56b19cc71556a1c

Rahmawati, T., Wilandari, Y., & Kartikasari, P. (2024). *Analisis perbandingan silhouette coefficient dan metode elbow pada pengelompokkan provinsi di indonesia berdasarkan indikator ipm dengan k-medoids 1,2,3*. *13*, 13–24. https://doi.org/10.14710/j.gauss.13.1.13-24

Rangkuti, Y. M., Nasution, S. A. P., Karo, I. M. K., & Fadhillah, W. N. (2023). *VISUALISASI PENYEBARAN COVID-19 DI KABUPATEN DELI SERDANG DENGAN SIG DAN K-MEANS*. Jejak Pustaka. https://ipusnas2.perpusnas.go.id/book/a97254ca-7361-4b8c-a61f-24376739fb9c

Refialy, L. P., Maitimu, H., & Pesulima, M. S. (2021). Perbaikan Kinerja Clustering K-Means pada Data Ekonomi Nelayan dengan Perhitungan Sum of Square Error (SSE) dan Optimasi nilai K cluster. *Techno.Com*, *20*(2), 321–329. https://doi.org/10.33633/tc.v20i2.4572

Rohmah Zaidah, A., Indira Septiarani, C., Sholikhatun Nisa, M., Yusuf, A., & Wahyudi, N. (2021). Komparasi Algoritma K-Means, K-Medoid, Agglomeartive Clustering Terhadap Genre Spotify. *Jurnal Ilmiah Ilmu Komputer*, *7*(1), 49–54. https://doi.org/10.35329/jiik.v7i1.186

Setiawan, W. (2021). *DATA MINING: TEORI DAN APLIKASI*. Rumah Cermelang Indonesia. https://ipusnas2.perpusnas.go.id/book/47802019-82e9-4543-8f88-b2ff614fb43b

Swasika, R., Mukodimah, S., Susanto, F., Muslihudin, M., & Ipnuwati, S. (2023). *IMPLEMENTASI DATA MINING (Clastering, Association, Prediction, Estimation, Classification)*. Adab. https://ipusnas2.perpusnas.go.id/book/c6fb2403-1432-41cc-b3f0-8ade254cd10b

Vania, P., & Nurina Sari, B. (2023). Perbandingan Metode Elbow dan Silhouette untuk Penentuan Jumlah Klaster yang Optimal pada Clustering Produksi Padi menggunakan Algoritma K-Means. *Jurnal Ilmiah Wahana Pendidikan*, *9*(21), 547–558. https://doi.org/10.5281/zenodo.10081332

Wala, J., Herman, & Umar, R. (2024). Implementasi K-Means Clustering pada

Pengelompokan Pasien Penyakit Jantung. *JISKA (Jurnal Informatika Sunan Kalijaga)*, *9*(3), 205–216.

WHO. (2024). *HIV dan AIDS*. World Health Organization. Diakses dari: https://www.who.int/news-room/fact-sheets/detail/hiv-aids

Zuliansyah, R. A. (2024). *189 Ibu Rumah Tangga di Banten Terpapar HIV, Paling Banyak di Tangerang Raya*. Tangerang Media. Diakses dari: https://www.tangerangnews.com/banten/read/51816/189-Ibu-Rumah-Tangga-di-Banten-Terpapar-HIV-Paling-Banyak-di-Tangerang-Raya

DOI: https://doi.org/10.17509/e.v24i2.84914
p- ISSN 2528-1410 e- ISSN 2527-8045