

EVALUASI K-MEANS DAN FUZZY-C-MEANS CLUSTERING ANALISIS DATA HIV DI TINGKAT PROVINSI

Gabriel Arnold Budianto

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara,
Jln. Letjen S. Parman No. 1, Jakarta, 11440, Indonesia
Email: gabriel.535210061@stu.untar.ac.id

ABSTRAK

HIV atau *Human Immunodeficiency Virus* adalah suatu penyakit yang berbahaya karena merusak imun tubuh dengan menginfeksi darah dan sel-sel didalam tubuh. Penyakit ini bisa mematikan dan harus diwaspadai karena penyebaran penyakit ini sangat tumbuh dengan cepat. Penyakit HIV cepat menular dengan cara seks bebas atau berhubungan dengan orang yang sudah terkena HIV melalui air liur atau seorang perempuan menyusui, bahkan bisa melalui kontaminasi alat Kesehatan seperti transfusi darah. Di provinsi Jawa Timur Surabaya kasus HIV sangat tinggi, maka dalam kasus ini diperlukan analisis data tentang penyebaran penyakit kasus HIV di provinsi Jawa Timur kota Surabaya berikut data pengamanan yang sesuai standar Kesehatan nasional Indonesia di berbagai puskesmas. Pada analisis data ini diperoleh dari website katalog.data.co.id. Proses analisis menggunakan metode *clustering* K-Means dan *Fuzzy-C-Means* untuk mengevaluasi dan mendata berapa banyak orang yang terjangkit HIV dan yang mendapatkan pelayanan berstandar kesehatan nasional Indonesia di setiap puskesmas yang ada di Surabaya. Hasil pengelompokkan evaluasi ini dapat digunakan untuk memetakan penyebaran penyakit HIV untuk berupaya tidak tersebar luas dalam bentuk penanganannya.

Kata kunci— *Evaluasi, Clustering, HIV, K-means, Fuzzy-C-means*

ABSTRACT

HIV or Human Immunodeficiency Virus is a dangerous disease because it damages the body's immunity by infecting the blood and cells in the body. This disease can be deadly and should be aware because the spread of this disease is very fast growing. HIV disease is quickly transmitted by way of free sex or in contact with people who have HIV through saliva or a breastfeeding woman, can even be through contamination of medical devices such as blood transfusions. In the province of East Java Surabaya HIV cases are very high, so in this case it is necessary to analyze data on the spread of HIV cases in the province of East Java Surabaya city along with treatment data according to Indonesian national health standards in various health centers. In this data analysis obtained from the website katalog.data.co.id. The analysis process uses K-Means and Fuzzy-C-Means clustering methods to evaluate and record how many people are infected with HIV and who get Indonesian national health standard services at each health center in Surabaya. The results of this evaluation clustering can be used to map the spread of HIV disease to try not to spread widely in the form of treatment.

Keywords - *Evaluation, Clustering, HIV, K-means, Fuzzy-C-means*

1.PENDAHULUAN

Human Immunodeficiency Virus (HIV) adalah virus yang dapat merusak kekebalan sistem tubuh terutama imun dan sel darah putih mengalami kerusakan pada bagian limfosit pada tubuh. Kasus ini dapat berakibat turunnya sistem kekebalan tubuh penderita HIV. Banyak dari beberapa orang mampu hidup berdampingan dengan tidak tampak seorang terkena penyakit HIV dan ada kasus

Begitu meningkatnya kasus penyakit HIV di Indonesia telah banyak terjadi di wilayah kota. HIV meular dengan cara berhubungan seksual atau seks bebas dalam pergaulan, transfuse darah dan air liur yang terkena HIV bahkan seorang perempuan menyusui terkena HIV [3]. Dengan demikian semua instansi terkait Kesehatan maupun keamanan nasional Indonesia harus bertindak signifikan dan akurat. *Machine learning* merupakan salah satu AI yang dapat digunakan untuk evaluasi penelitian berdasarkan penelitian terlebih dahulu yang dilakukan oleh Rut Indra Lita Sinaga, dkk., tentang “Pengelompokkan Jumlah Kasus Penyakit Aids berdasarkan Provinsi Menggunakan metode K-means” berdasarkan provinsi pada tahun 2016-2018 dengan metode K-Means terbagi menjadi 2 model yaitu Cluster 1 kategori rendah dan cluster 2 kategori tinggi [4]. Kemudian dengan metode Fuzzy-C-Means yang dapat melakukan penyebaran data pada seluruh cluster menggunakan FCM ini mencari tahu seberapa besar pengaruh yang dimiliki terhadap masing-masing data Penelitian evaluasi ini bertujuan untuk memberikan evaluasi terhadap semua warga provinsi Surabaya melalui puskesmas yang tersebar di kecamatan dan kelurahan dengan kasus HIV beresiko tinggi dan rendah dalam penanganan di kesehatan setempat dalam upaya penanggulangan penyebaran dan pencegahan secara luas..

	nama_sis	kekurahan_kode_sas	kecamatan	uang_lir	nama_kel	nama_peserta	jumlah_p	jumlah_p	deskrrips	jenis_per	total_per	jumlah_orang_terduga_menderita_hiv_yang_mendapatkan_pelayanan_kesehatan_sesuai_standar
1	Puskesmas Kel. Taml	1E+06	Kec. Ase SKDR	Surveil	3	0	0	0	Surveil	Surveil	0	
2	Puskesmas Kel. Sem	1E+06	Kec. Ber SKDR	Surveil	3	0	0	0	Surveil	Surveil	0	
3	Puskesmas Kel. Karu	1E+06	Kec. Ber SKDR	Surveil	3	0	0	0	Surveil	Surveil	0	
4	Puskesmas Kel. Laki	1E+06	Kec. Lak KLB	penega	3	0	0	0	penega	Penangs	1	
5	Puskesmas Kel. Lidal	1E+06	Kec. Lak KLB	penega	3	0	0	0	penega	Penangs	0	46
6	Puskesmas Kel. Sum	1E+06	Kec. Pak KLB	penega	3	1	0	0	penega	Penangs	1	
7	Puskesmas Kel. Jenu	1E+06	Kec. Lak KLB	penega	3	0	1	0	penega	Penangs	0	102
8	Puskesmas Kel. Lidal	1E+06	Kec. Lak KLB	penega	3	0	0	0	penega	Penangs	0	
9	Puskesmas Kel. Sum	1E+06	Kec. Lak KLB	penega	3	0	0	0	penega	Penangs	0	9
10	Puskesmas Kel. Bang	1E+06	Kec. Lak KLB	penega	3	0	0	0	penega	Penangs	0	
11	Puskesmas Kel. Sanl	1E+06	Kec. San KLB	penega	3	0	0	0	penega	Penangs	0	46
12	Puskesmas Kel. Karu	1E+06	Kec. Ber KLB	penega	3	0	0	0	penega	Penangs	0	
13	Puskesmas Kel. Rom	1E+06	Kec. Ber KLB	penega	3	0	0	0	penega	Penangs	0	
14	Puskesmas Kel. Tanc	1E+06	Kec. Tan SKDR	Surveil	3	0	0	0	Surveil	Surveil	0	
15	Puskesmas Kel. Kara	1E+06	Kec. Tan SKDR	Surveil	3	0	0	0	Surveil	Surveil	0	
16	Puskesmas Kel. Taml	1E+06	Kec. Ber KLB	penega	3	0	0	0	penega	Penangs	0	
17	Puskesmas Kel. Asar	1E+06	Kec. Ase SKDR	Surveil	3	0	0	0	Surveil	Surveil	0	
18	Puskesmas Kel. Tanc	1E+06	Kec. Suk KLB	penega	3	0	0	0	penega	Penangs	0	2
19	Puskesmas Kel. Sont	1E+06	Kec. Suk KLB	penega	3	0	0	0	penega	Penangs	0	
20	Puskesmas Kel. Puta	1E+06	Kec. Suk KLB	penega	3	0	0	0	penega	Penangs	0	
21	Puskesmas Kel. Sinc	1E+06	Kec. Suk KLB	penega	3	0	0	0	penega	Penangs	0	11
22	Puskesmas Kel. Sinc	1E+06	Kec. Suk KLB	penega	3	0	0	0	penega	Penangs	0	
23	Puskesmas Kel. Suk	1E+06	Kec. Suk KLB	penega	3	0	0	0	penega	Penangs	0	
24	Puskesmas Kel. Gert	1E+06	Kec. Ase SKDR	Surveil	3	0	0	0	Surveil	Surveil	0	
25	Puskesmas Kel. Bab	1E+06	Kec. Pak KLB	penega	3	0	0	0	penega	Penangs	0	64
26	Puskesmas Kel. Pakz	1E+06	Kec. Pak KLB	penega	3	0	0	0	penega	Penangs	0	
27	Puskesmas Kel. Beru	1E+06	Kec. Pak KLB	penega	3	0	0	0	penega	Penangs	0	
28	Puskesmas Kel. Taml	1E+06	Kec. Ber KLB	penega	3	0	0	0	penega	Penangs	0	
29	Puskesmas Kel. Lont	1E+06	Kec. San KLB	penega	3	0	0	0	penega	Penangs	0	
30	Puskesmas Kel. Bidi	1E+06	Kec. San KLB	penega	3	0	0	0	penega	Penangs	0	27
31	Puskesmas Kel. Bidi	1E+06	Kec. San KLB	penega	3	0	0	0	penega	Penangs	0	
32	Puskesmas Kel. Bidi	1E+06	Kec. San KLB	penega	3	0	0	0	penega	Penangs	0	10
33	Puskesmas Kel. Peng	1E+06	Kec. Ger KLB	penega	3	0	0	0	penega	Penangs	0	
34	Puskesmas Kel. Gert	1E+06	Kec. Ger KLB	penega	3	0	0	0	penega	Penangs	0	
35	Puskesmas Kel. Kapi	1E+06	Kec. Ger KLB	penega	3	0	0	0	penega	Penangs	0	
36	Puskesmas Kel. Kara	1E+06	Kec. Ger KLB	penega	3	0	0	0	penega	Penangs	0	30
37	Puskesmas Kel. Emb	1E+06	Kec. Ger KLB	penega	3	0	0	0	penega	Penangs	0	
38	Puskesmas Kel. Ked	1E+06	Kec. Teg KLB	penega	3	1	0	0	penega	Penangs	1	147
39	Puskesmas Kel. Tega	1E+06	Kec. Teg KLB	penega	3	0	0	0	penega	Penangs	0	
40	Puskesmas Kel. Dr. S	1E+06	Kec. Teg KLB	penega	3	0	0	0	penega	Penangs	0	15
41	Puskesmas Kel. Won	1E+06	Kec. Teg KLB	penega	3	0	0	0	penega	Penangs	0	
42	Puskesmas Kel. Kepi	1E+06	Kec. Teg KLB	penega	3	0	0	0	penega	Penangs	0	
43	Puskesmas Kel. Alur	1E+06	Kec. Pak KLB	penega	3	0	0	0	penega	Penangs	0	38
44	Puskesmas Kel. Teml	1E+06	Kec. But KLB	penega	3	0	0	0	penega	Penangs	0	
45	Puskesmas Kel. Jeps	1E+06	Kec. But KLB	penega	3	0	0	0	penega	Penangs	0	
46	Puskesmas Kel. Guri	1E+06	Kec. But KLB	penega	3	0	0	0	penega	Penangs	0	72
47	Puskesmas Kel. Bubi	1E+06	Kec. But KLB	penega	3	0	0	0	penega	Penangs	0	
48	Puskesmas Kel. Taml	1E+06	Kec. Sim KLB	penega	3	0	0	0	penega	Penangs	0	45
49	Puskesmas Kel. Kapi	1E+06	Kec. Sim KLB	penega	3	0	0	0	penega	Penangs	0	
50	Puskesmas Kel. Bim	1E+06	Kec. Sim KLB	penega	3	0	0	0	penega	Penangs	0	

2.2 Clustering Data

Dalam penelitian ini menggunakan model clustering untuk mengevaluasi data yang dilakukan dengan cara unsupervised learning dalam machine learning menggunakan dataset csv dari Katalog.data.co.id. Metode Clustering ini dengan dua algoritma machine learning yaitu K-means dan Fuzzy-c-Means beserta silhouette dan davies bouildin index. Pengelompokan data diperlukan karena data yang belum diolah tidak mudah dilakukan analisis tujuannya untuk memahami data.

2.2.1 Metode K-Means

Algoritma K-Means unruk melakukan clustering melakukan penentuan nilai K atau cluster yang akan dibuat dan secara acak akan mendistribusikan data cluster kemudian menghitung rata – rata dari data K-means metode sederhana untuk analisis clustering untuk menentukan membagi entitas ke-n menjadi kelompok yang disebut cluster. Setiap cluster terbentuk meningkatkan kriteria fungsi. Euclidean digunakan sebagai jarak dalam metode K-means dengan rumus:

titik $a = (a_1, a_2, \dots, a_K)$ dan titik $b = (b_1, b_2, \dots, b_n)$ dapat dihitung dengan Rumus (1). $d(b_i, a_t) = \sqrt{\sum (b_{ij} - a_{tj})^2}$ (1) dimana: d = jarak antara nilai data dan nilai pusat cluster b_i = nilai data

2.2.2 Metode Fuzzy-C-Means

Menurut Bezdek (dalam Afiyah, 2014), *analisis clustering fuzzy* menggunakan tingkat keanggotaan dan *cluster fuzzy* sebagai pembobot untuk pengelompokan. Metode ini merupakan evolusi dari metode pembagian data yang menggunakan *Fuzzy clustering* memiliki keunggulan utama, yaitu dapat memberikan hasil pengelompokan untuk objek yang tersebar tidak teratur. Ini disebabkan oleh kemungkinan bahwa suatu titik data memiliki sifat atau karakteristik dari suatu *cluster* lain, sehingga kecenderungan titik data terhadap suatu cluster. Bezdek (1981) merumuskan masalah *fuzzy clustering* secara matematis sebagai optimasi kendala (Efiyah, 2014). Sebelum melakukan clustering fuzzy, beberapa hal harus diketahui (Kusumadewi dkk., 2006): 1. *Cluster Fuzzy* dengan rumus berikut dapat dilihat pada gambar 2.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w \cdot x_{ij})}{\sum_{i=1}^n (\mu_{ik})^w}$$

Gambar 2. rumus Fuzzy C Means

V_{kj} = pusat cluster untuk atribut j

N_{ik} = derajat anggota untuk data sampel ke I dan k

X_{ij} data ke I dan ke j

2.2.3 Evaluasi Clustering Silhouette

Metode evaluasi *silhouette* dikembangkan oleh Rousseeuw. Nilai silhouette terletak pada rentang 1 banding 1 kemudian *clustering* yang baik memiliki nilai *silhouette* positif. *Silhouette* dihitung menggunakan formula dapat dilihat pada gambar 3.

Evaluasi Clustering: Silhouette

Nilai Silhouette	Interpretasi
0.71 - 1.00	Struktur yang dihasilkan kuat
0.51 - 0.70	Struktur yang dihasilkan baik
0.26 - 0.50	Struktur yang dihasilkan lemah
≤ 0.25	Tidak terstruktur

Gambar 3 Evaluasi Clustering Silhouette

2.2.4 Evaluasi *davies boildin index*

Evaluasi pada metode *fuzzy c means* untuk mengetahui hasil baik *clustering* yang dihasilkan menghitung kuantitas dari turunan data yang diperoleh untuk memaksimalkan jarak antar cluster dengan rumus sebagai dapat dilihat pada gambar 4.

$$R_i = \max_{j=1, \dots, k, i \neq j} R_{ij} \quad R_{ij} = \frac{\text{var}(C_i) + \text{var}(C_j)}{\|c_i - c_j\|}$$

Gambar 4 Rumus *davies Boildin Index*

Experiment	Davies Bouldin Index Value					
	Simple K-Means			K-Means with centroid based min SSE		
	k=2	k=3	k=4	k=2	k=3	k=4
1	0.4296	0.0524	0.0406	0.1366	0.0493	0.0372
2	0.4296	0.0491	0.0358	0.2392	0.0540	0.0409
3	0.4296	0.0524	0.0322	0.2392	0.0505	0.0373
4	0.4296	0.0661	0.0367	0.2392	0.0493	0.0409
5	0.2845	0.0493	0.0391	0.2392	0.0493	0.0372
6	0.2845	0.0679	0.0515	0.2392	0.0493	0.0319
7	0.2845	0.0493	0.0320	0.2392	0.0493	0.0373
8	0.4296	0.0576	0.0407	0.2392	0.0491	0.0409
9	0.4296	0.0493	0.0365	0.1366	0.0493	0.0365
10	0.4296	0.0661	0.0407	0.2392	0.0493	0.0406
Average	0.38607	0.05595	0.03858	0.21868	0.04987	0.03807

Gambar 5 Evaluasi Davies Boildin index

3. HASIL DAN PEMBAHASAN

3.1. *Performa K-Means*

Pengelompokan data dengan metode K-Means terjadi dimana semua jumlah data per kategori sudah dikelompokkan berdasarkan jenis dan jumlah total data yang ada pada gambar 6. Kemudian statistik data elbow yang muncul dapat dilihat pada gambar 7.

```
[ ] Karakteristik rata-rata setiap cluster:
[+] kode_sarana_kesehatan lama_penanganan \
Cluster
0.0 1.033304e+06 3.0
1.0 1.033316e+06 3.0
2.0 1.033326e+06 3.0

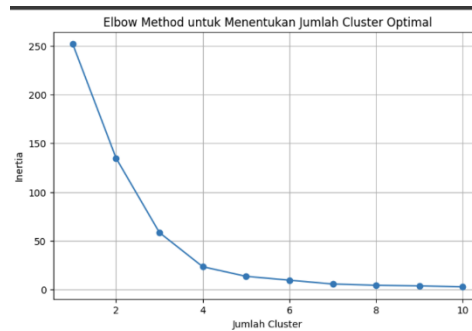
jumlah_penduduk_terjangkit_laki2 \
Cluster
0.0 1.0
1.0 0.0
2.0 0.0

jumlah_penduduk_terjangkit_perempuan total_penduduk_terjangkit \
Cluster
0.0 0.0 1.0
1.0 0.0 0.0
2.0 1.0 1.0

jumlah_orang_terduga_menderita_hiv_yang_mendapatkan_pelayanan_kesehatan_sesuai_standar \
Cluster
0.0 83.400000
1.0 56.631579
2.0 1.000000

Cluster
Cluster
0.0 0.0
1.0 1.0
2.0 2.0
```

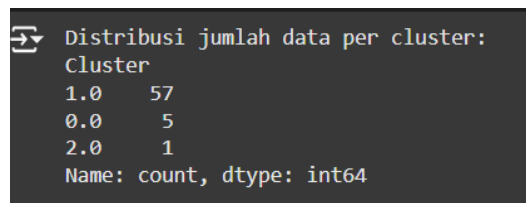
Gambar 6 Pengelompokan *K-Means*



Gambar 7 elbow K-Means

3.2. Performa Fuzzy C means

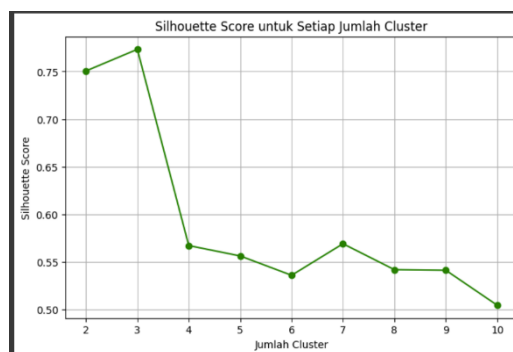
Parameter yang digunakan pada model ini yaitu random state yang merupakan pembuat bilangan acak dalam clustering. Model ini menetapkan nilai random_state dimulai dari angka 0. Hasil dari metode FCM dapat dilihat pada Gambar 4.



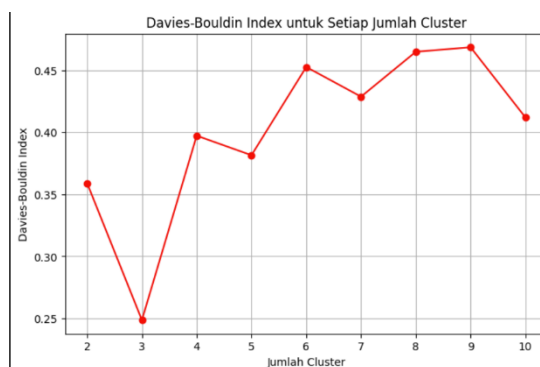
Gambar 8 hasil clustering Fuzzy C Means

3.3. Evaluasi menggunakan Silhouette dan Davies Boildin Index

Masing-masing metode ini melampirkan hasil *elbow* yang ada kemudian setelah di *clustering* data akan muncul Kembali dengan data yang sebenarnya dapat dilihat pada gambar 9 metode *silhouette* dan gambar 10 *davies boildin*.



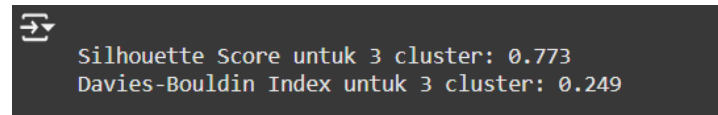
Gambar 9 Silhouette score



Gambar 10 Davies Boildin Index

3.4. Perbandingan Performa Algoritma

Metode clustering yang dilakukan dibandingkan akurasi dan evaluasi agar dapat terlihat metode mana yang memiliki performa terbaik. Perbandingan evaluasi clustering ini dapat dilihat pada Gambar 11.



Gambar 11. perbandingan akurasi clustering

4. KESIMPULAN

Perbandingan evaluasi machine learning yang telah diuji coba menghasilkan nilai akurasi yang cukup akurat yaitu *K-Means silhouette* dengan cepat 3 cluster dengan akurasi 0.773 kemudian *Fuzzy C-Means Davies Boildin index* dengan cepat 3 cluster dengan akurasi 0.249. Komparasi Algoritma *Machine Learning Clustering Fuzzy-C-Means* lebih akurat daripada *K-Means* saat melakukan clustering analisis data HIV dengan penanganan berstandar nasional Indonesia di Provinsi Jatim Surabaya. Hasil penelitian ini perlu di uji pada data yang lebih besar untuk lebih detail mengetahui hasil evaluasi.

DAFTAR PUSTAKA

- [1] A. W. P. Samudra, R. A. Susanto, A. R. Putra, F. I. Kurniadi, dan B. Juarto, "Klasifikasi HIV AIDS dengan Aplikasi Rapid Miner," *Sist. Komput. dan Kecerdasan Buatan*, no. 1, hal. 15–19, 2022.
- [2] R. Alhamidi, "Membaca Grafik Kasus HIV/AIDS di Jabar: Naik 3 Tahun Terakhir!," *detikjabar*, 2022. <https://www.detik.com/jabar/berita/d6262009/membaca-grafik-kasus-hiv-aids-dijabar-naik-3-tahun-terakhir> (diakses 18 Januari 2023).
- [3] A. A. Lesmana et al., "Implementasi Algoritma K-Means Untuk Clustering Penyakit Hiv / Aids Di Indonesia Implementation of Kmeans Algorithm for Clustering of Hiv / Aids Disease in Indonesia," *eProceeding Eng.*, vol. 6, no. 2, hal. 5564–5580, 2019.
- [4] R. I. L. Sinaga, W. Saputra, dan H. Qurniawan, "Pengelompokan Jumlah Kasus Penyakit Aids Berdasarkan Provinsi Menggunakan Metode KMeans," *Kesatria Jurnal Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 2, no. 2, hal. 99–107, 2021.
- [5] dosbing.id, "Proses tahapan data mining," *dosbing.id*, 2023. <https://dosbing.id/2023/04/21/proses-tahapandata-mining/> (diakses 4 Februari 2023).
- [6] A. Y. E. Dodu, D. W. Nugraha, dan M. A. Ayyub, "Penerapan Data Mining Untuk Mendeteksi Tingkatan Stadium Penyakit *Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome* (Hiv/Aids) Menggunakan Algoritma Naive Bayes Classifier (Studi Kasus Pada Rumah Sakit Umum Daerah (Rsud) Anutapura Palu)," *Sci. Comput. Sci. Informatics J.*, vol. 1, no. 1, hal. 33, 2019, doi: 10.22487/j26204118.2018.v1.i1.11900.
- [7] V. Herlinda dan D. Darwis, "Analisis Clustering Untuk Recredesialing Fasilitas Kesehatan Menggunakan Metode Fuzzy C-Means," *Teknol. dan Sist. Inf.*, vol. 2, no. 2, hal. 94–99, 2021, [Daring]. Tersedia pada: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [8] H. Noor, A. Dharmawati, dan T. W. Qur'ana, "Penerapan Algoritma K-Means Clustering Analysis Pada Kasus Penderita Hiv/Aids (Studi Kasus Kabupaten Banjar)," *Technol. J. Ilm.*, vol. 12, no. 2, hal. 72, 2021, doi: 10.31602/tji.v12i2.4573.
- [9] D. Y. Nugroho, A. Wibowo, F. Kesehatan, dan M. Universitas, "Analisis Cluster K - Means Kabupaten / Kota Di Jawa Timur Berdasarkan Level Epidemik Hiv Cluster K - Means Analysis of The HIV Epidemic Level In," *Biometrika dan Kependud.*, vol. 8, no. 2, hal. 108–117, 2019, doi: 10.20473/jbk.v8i2.2019.108.
- [10] opendata.jabarprov.go.id, "Jumlah Kasus HIV Berdasarkan Kelompok Umur di Jawa Barat," *opendata.jabarprov.go.id*, 2021. <https://opendata.jabarprov.go.id/id/dataset/jumlah-kasus-hiv-berdasarkan-kelompok-umur-dijawa-barat> (diakses 18 Januari 2023).

- [11] D. Latifatul Fajri, "Studi Pustaka Adalah Referensi Penelitian, Ini Penjelasan Lengkapnya," katadata, 2022. <https://katadata.co.id/agung/berita/62e773e3da762/studi-pustaka-adalah-referensi-penelitian-inipenjelasan-lengkapnya> (diakses 27 Januari 2023).
- [12] E. R. Wardhani, "Manajemen Perubahan Dalam Implementasi Reformasi Birokrasi Internal," 2018. [Daring]. Tersedia pada: http://repository.upi.edu/44420/7/T_ADPEN_1608369_Chapter3.pdf
- [13] H. Noor, A. Dharmawati, and T. W. Qur'ana, "Penerapan Algoritma K-Means Clustering Analysis Pada Kasus Penderita Hiv/Aids (Studi Kasus Kabupaten Banjar)," Technol. J. Ilm., vol. 12, no. 2, p. 72, 2021, doi: 10.31602/tji.v12i2.4573.
- [14] D. Rohmatullailah and D. Fikriyah, "Faktor Risiko Kejadian HIV Pada Kelompok Usia Produktif di Indonesia," J. Biostat. Kependudukan, dan Inform. Kesehatan, vol. 2, no. 1, p. 45, 2021, doi: 10.51181/bikfokes.v2i1.4652.
- [15] R. I. L. Sinaga, W. Saputra, and H. Qurniawan, "Pengelompokan Jumlah Kasus Penyakit Aids Berdasarkan Provinsi Menggunakan Metode K-Means," J. Penerapan Sist. Inf. (Komputer Manajemen), vol. 2, no. 2, pp. 99–107, 2021.
- [16] D. Y. Nugroho, A. Wibowo, F. Kesehatan, and M. Universitas, "Analisis Cluster K - Means Kabupaten / Kota Di Jawa Timur Berdasarkan Level Epidemik HIV Cluster K - Means Analysis of The HIV Epidemic Level In," vol. 8, no. 2, pp. 108–117, 2019, doi: 10.20473/jbk.v8i2.2019.108.
- [17] Z. Khan, "HIV-Infected People Data," vol. 100, no. 13, pp. 5010–5018, 2022.
- [18] N. A. Mufid, D. Risqi, N. Hanum, and A. H. Sidiq, "Clustering Kabupaten / Kota di Jawa Tengah Tahun 2022 berdasarkan Jumlah Kasus Kemunculan Penyakit dengan Algoritma K-Means," no. 1, pp. 1–14, 2023.
- [19] F. Novianti, Y. R. Aisyah Yasmin, and D. C. R. Novitasari, "Penerapan Algoritma Fuzzy C-Means (FCM) dalam Pengelompokan Provinsi di Indonesia berdasarkan Indikator Penyakit Menular Manusia," JUMANJI (Jurnal Masy. Inform. Unjani), vol. 6, no. 1, p. 23, 2022, doi: 10.26874/jumanji.v6i1.103.
- [20] T. P. Sari, A. L. Hananto, E. Novalia, T. Tukino, and S. S. Hilabi, "Implementasi Algoritma K-Means dalam Analisis Klasterisasi Penyebaran Penyakit Hiv/Aids," Infotek J. Inform. dan Teknol., vol. 6, no. 1, pp. 104–114, 2023, doi: 10.29408/jit.v6i1.7423.
- [21] E. M. Fitri, R. R. Suryono, and A. Wantoro, "Klasterisasi Data Penjualan Berdasarkan Wilayah Menggunakan Metode K-Means Pada Pt Xyz," J. Komputasi, vol. 11, no. 2, pp. 157–168, 2023, doi: 10.23960/komputasi.v11i2.12582.
- [22] W. Rianti and E. Harahap, "Pengolahan Data Hasil Penjualan Online Menggunakan Aplikasi Microsoft Excel," Mat. J. Teor. dan Terap. Mat., vol. 20, no. 2, pp. 69–76, 2021, [Online]. Available: <https://journals.unisba.ac.id/index.php/matematika/article/view/1553>
- [23] P. S. Rosiana, A. A. Mohsa, and Y. Umaidah, "Implementasi K-Means Dalam Pengelompokan Penyebaran Penyakit Dbd Di Jawa Barat," J. Inform. dan Tek. Elektro Terap., vol. 11, no. 3, pp. 782–788, 2023, doi: 10.23960/jitet.v11i3.3344.
- [24] D. I. Yunistya, R. Goejantoro, F. Deny, and T. Amijaya, "The Application Of K-Harmonic Means Method in District/City Grouping (Case Study: Poverty in Kalimantan Island in 2020) Penerapan Metode K-Harmonic Means Dalam Pengelompokan Kabupaten/Kota (Studi Kasus: Kemiskinan di Pulau Kalimantan Tahun 2020)," J. Mat. Stat. dan Komputasi, vol. 19, no. 1, pp. 51–64, 2022, doi: 10.20956/j.v19i1.21116.
- [25] A.- Shoolihah, M. T. Furqon, and A. W. Widodo, "Implementasi Metode *Improved K-Means* untuk Mengelompokkan Titik Panas Bumi," vol. 1, no. 11, pp. 1270–1276, 2019.