Seminar Nasional Informatika (SENATIKA)
Prosiding Senatika 2025 **QRSBN:62-0151-02942-8**

Analisis Penyakit Menular dan Tidak Menular Berdasarkan Wilayah Tempat Tinggal Menggunakan Metode *K-Means* dan *Fuzzy C-Means* (Studi Kasus: Puskesmas Bayung Lencir)

Marsa Juleha[1], Rike Limia Budiarti[2], Novhirtamely Kahar [3]
[1,2,3]Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Nurdin Hamzah, Jambi, Indonesia
e-mail: [1]marsajeha@gmail.com, [2]rikelimia@gmail.com, [3]novmely@gmail.com

Abstrak

Penyakit menular dan tidak menular masih menjadi tantangan utama dalam pelayanan kesehatan masyarakat. Penelitian ini dilakukan di Puskesmas Bayung Lencir dengan tujuan untuk melakukan analisis risiko penyakit berdasarkan data jumlah kasus penyakit Demam Berdarah Dengue (DBD), Tuberkulosis (TBC), Hipertensi, dan Diare. Metode yang digunakan adalah K-Means dan Fuzzy C-Means, dua teknik clustering populer dalam data mining yang akan dibandingkan dari segi efektivitas dalam mengelompokkn data menjadi dua klaster, yaitu klaster risiko penyakit cukup rendah dan klaster risiko penyakit cukup tinggi. Data yang digunakan adalah data tahunan yang telah melalui prosaes normalisasi. Hasil analisis menunjukkan bahwa Fuzzy C-Means (FCM) memiliki akurasi lebih baik dibandingkan K- Means, dengan nilai validasi Silhouette Coefficient mencapai 84%, sedangkan K-Means hanya mencapai sekitar 80%. Hal ini disebabkan oleh kemampuan FCM dalam menangkap ketidakpastian data dengan memberikan derajat keanggotaan pada tiap klaster, sehingga lebih adaptif terhadap data yang saling tumpang tindih. Sebaliknya, K-Means menghasilkan batas klaster yang tegas namun kurang fleksibel.

Kata Kunci: *Clustering, Fuzzy C-Means, K-Means, penyakit menular, penyakit tidak menular, Data mining*

Abstract

Infectious and non-communicable diseases remain a major challenge in public health services. This study was conducted at the Bayung Lencir Community Health Center with the aim of analyzing disease risks based on the number of cases of Dengue Hemorrhagic Fever (DHF), Tuberculosis (TB), Hypertension, and Diarrhea. The methods used are K-Means and Fuzzy C-Means (FCM), two popular clustering techniques in data mining, which are compared in terms of their effectiveness in grouping data into two clusters: relatively low disease risk and relatively high disease risk. The dataset used is annual health records that have undergone a normalization process. The results of the analysis show that Fuzzy C-Means (FCM) achieves higher accuracy than K-Means, with a Silhouette Coefficient validation value of 84%, compared to 80% for K-Means. This superiority is due to FCM's ability to capture data uncertainty by assigning membership degrees across clusters, making it more adaptive to overlapping data. In contrast, K-Means produces rigid cluster boundaries, which are less flexible when handling ambiguous cases.

Keywords: *Data mining, Clustering, Fuzzy C-Means, infectious diseases, K-Means, non-infectious diseases*

1. Pendahuluan

Masyarakat dan lingkungan memiliki keterkaitan erat dalam memengaruhi kondisi kesehatan wilayah tempat tinggal. Kebersihan lingkungan, pola pembuangan sampah, sanitasi air, serta pencemaran udara dan tanah dapat meningkatkan risiko penyakit menular maupun tidak menular. Puskesmas Bayung Lencir masih menggunakan pencatatan manual yang hanya mengurutkan penyakit terbanyak, tanpa memperhatikan atribut seperti alamat, jenis kelamin, dan jenis peserta. Kondisi ini membuat analisis pola penyebaran penyakit menjadi terbatas, padahal informasi tersebut penting untuk kebijakan kesehatan masyarakat [1], [2], [3], [4], [5].

Beberapa penelitian sebelumnya telah menerapkan metode clustering untuk pengelompokan data penyakit. Penelitian oleh Puspitasari [6] menggunakan algoritma K-Means untuk mengelompokkan data penyakit menular berdasarkan wilayah, namun hasilnya masih

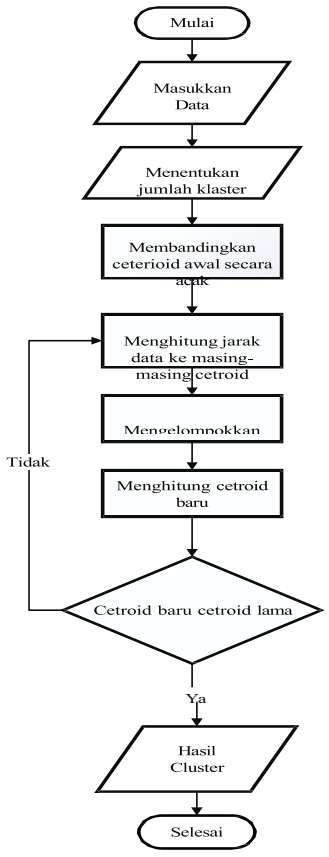
dipengaruhi oleh nilai inisialisasi awal yang acak sehingga akurasi pengelompokan kurang optimal. Sementara itu, penelitian oleh Fadilah dkk. [7] menggunakan algoritma Fuzzy C-Means (FCM) untuk pengelompokan daerah rawan penyakit demam berdarah dan memperoleh hasil yang lebih fleksibel, karena setiap data dapat memiliki derajat keanggotaan dalam beberapa klaster. Namun, penelitian tersebut belum melakukan perbandingan langsung antara kedua algoritma pada data penyakit dengan faktor lingkungan yang beragam.

Berdasarkan hal tersebut, penelitian ini memiliki gap yaitu belum adanya analisis perbandingan antara algoritma K-Means dan Fuzzy C-Means untuk pengelompokan penyakit menular berdasarkan data lingkungan. Kedua algoritma ini penting karena memiliki karakteristik berbeda: K-Means bekerja dengan pendekatan keras (hard clustering) di mana satu data hanya dapat masuk ke satu klaster, sedangkan Fuzzy C-Means menggunakan pendekatan lembut (soft clustering) yang memberikan derajat keanggotaan. Dengan membandingkan kedua algoritma tersebut, dapat diketahui metode mana yang lebih efektif dan sesuai dengan karakteristik data penyakit di wilayah Puskesmas Bayung Lencir[8].

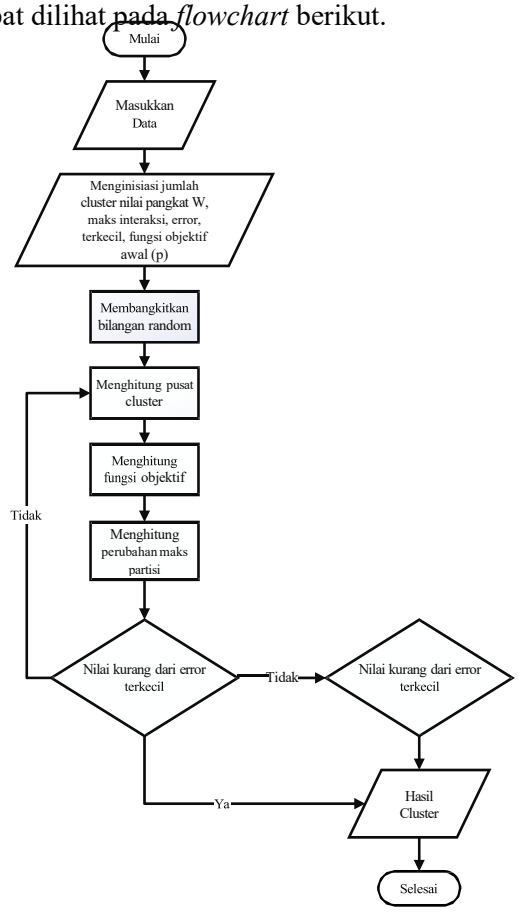
Hasil klasterisasi dapat dimanfaatkan oleh Puskesmas untuk menentukan wilayah berisiko tinggi, mendeteksi kelompok rentan, serta menyusun strategi pencegahan dan intervensi [9], [10]. Misalnya, wilayah dengan sanitasi buruk dapat difokuskan untuk penyuluhan kebersihan, penyediaan air bersih, atau pengadaan obat [11]. Implementasi metode ini dapat dilakukan dengan bahasa pemrograman *Python* karena kemudahan, pustaka yang kuat, serta dukungan komunitas yang luas [12], [13], [14], [15]. Dengan demikian, penelitian berjudul *"*Analisis Penyakit Menular dan Tidak Menular Berdasarkan Wilayah Tempat Tinggal Menggunakan Metode *K- Means* dan *Fuzzy C-Means"* diharapkan mampu menghasilkan analisis yang bermanfaat untuk mendukung pengambilan keputusan kesehatan di Puskesmas Bayung Lencir.

2. Metode Penelitian

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan *data mining*, dengan menggunakan 2 metode klasterisasi yaitu *K-Means* dan *Fuzzy C-Means*. Tahapan metode *K-Means* dan *Fuzzy C-Means* dapat dilihat pada *flowchart* berikut.



Gambar 1. *Flowchart* Metode *K-Means*



Gambar 2. *Flowchart* Metode *Fuzzy C-Means*

Alur algoritma *K-Means* pada gambar 1 *flowchart* di atas juga dapat dijelaskan dengan cara

sebagai berikut:

1. Tentukan nilai k sebagai jumlah klaster yang ingin dibentuk.
2. Inisialisasi k pusat cluster ini ari dilakukan dengan berbagai cara, namun yang paling sering dilakukan adalah dengan cara random yang diambil dari data yang ada.
3. Menghitung jarak setiap data input terhadap masing – masing centroid menggunakan rumus jarak Euclidean (Euclidean Distance) hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. Berikut adalah persamaan Euclidian Distance: $De = \sqrt{(xi - si)}\ 2(yi - ti)$ 2 dimana : De adalah Euclidean Distance i adalah banyaknya objek, (x,y) merupakan koordinat object dan (s,t) merupakan koordinat centroid.
4. Mengklasifikasikan setiap data berdasarkan kedekatannya dengan centroid (jarak terkecil).
5. Memperbarui nilai centroid. Nilai centroid baru di peroleh dari rata-rata cluster yang bersangkutan dengan menggunakan rumus: $vij = 1\ Ni \sum Xkj\ Ni\ k$=0 dimana: vij adalah centroid/ rata-rata cluster ke-i untuk variable ke-j Ni adalah jumlah data yang menjadi anggota cluster ke-i i, k adalah indeks dari cluster j adalah indeks dari ariable xkj adalah nilai data ke-k yang ada di dalam cluster tersebut untuk variable ke-j[16], [17].

Alur algoritma *Fuzzy C-Means* pada gambar 2 *flowchart* di atas juga dapat dijelaskan dengan cara sebagai berikut:

1. Input data yang akan di cluster X, dalam bentuk matriks berukuran $n \times m$ (n = jumlah sampel data, m = atribut setiap data). Xij= data sampel ke-i (i = 1,2,3,…,n), atribut ke-j (j = 1,2,…,m).
2. Parameter yang dibutuhkan, yaitu: a. Jumlah klaster I b. Pangkat/bobot (w) c. Maksimum iterasi (MaxIter) d. Error terkecil (ε) e. Fungsi objektif ($Po = 0$) f. Iterasi awal (t = 1)
3. Bentuk bilangan random μik (i = 1,2,…,n dan k = 1,2,…,c) yang merupakan elemenelemen matriks partisi awal U. kemudian, hitung jumlah setiap kolom dengan persamaan (1). $Qi = \sum$ µik $c\ k$=1 (1) Dimana : μik = derajat eanggotaan Qi = Jumlah nilai derajat keanggotaan perkolom = 1 dengan i=1,2,…,n;Tentukan nilai matriks partisi awalmenggunakan persamaan (2). $\mu ik = \mu ik\ Qi$ (2)
4. Hitung pusat cluster ke-k : Vkj, dimana k = 1,2,…,c dan j = 1,2,..,n menggunakan persamaan (3). $Vkj = \sum ((\mu ik)\ w{*}Xij)\ n\ i$=1 $\sum n\ (\mu ik)\ i$=1 (3) Dimana : V = pusat klaster Xi = parameter ke-i [18], [19].
5. Hitung fungsi objektif pada iterasi ke-t ( ) dengan persamaan (4). $Pt = \sum \sum ([\sum (Xij - Vkj)\ m\ 2\ j$=1]$(\mu ik)\ w)\ c\ k$=1 $n\ i$=1 (4) Dimana Pt adalah nilai fungsi objektif iterasi ke-t .
6. Hitung perubahan matriks partisi menggunakan persamaan (5). $\mu ik = [\sum (Xij{-}Vkj)\ 2]\ m\ j$=1 −1 w−1 $\sum [\sum (Xij{-}Vij)\ 2]$ −1 $m\ w$−1 j=1 $c\ k$=1 (5) Dimana i = 1,2,…,n dan k = 1,2,…,c.
7. Cek kondisi berhenti, dengan ketentuan yaitu : a. $(Pt - (Pt - 1)| < \varepsilon)$ atau $(t > MaxIter)$ maka berhenti b. jika tidak : $t = t + 1$ maka ulangi langkah ke-4. [20]

2.1 Tahap Penelitian

1. Pengumpulan Data
   Data diperoleh dari Puskesmas Bayung Lencir tahun 2025, meliputi empat jenis penyakit: Demam Berdarah Dengue (DBD), Tuberkulosis (TBC), Hipertensi, dan Diare.
2. Pra-pemrosesan Data (Preprocessing)
   Tahapan ini mencakup pembersihan data, penanganan data hilang, dan normalisasi agar data berada pada rentang [0,1] menggunakan metode *Min-Max Normalization*.
3. Penentuan Parameter Algoritma
   a. Jumlah klaster (k) = 2 (kategori risiko rendah dan tinggi)
   b. Pembobot (*fuzziness*) FCM = 2
   c. *Error threshold* = 0.0001
   d. Iterasi maksimum = 100
4. Penentuan Parameter Algoritma
   a. Jumlah klaster (k) = 2 (kategori risiko rendah dan tinggi)
   b. Pembobot (*fuzziness*) FCM = 2
   c. Error threshold = 0.0001
   d. Iterasi maksimum = 100
5. PenerapanAlgoritma
   a. *K-Means:* menghitung jarak Euclidean dari setiap data ke centroid, mengelompokkan data berdasarkan jarak terkecil, dan memperbarui centroid hingga konvergen.
   b. *Fuzzy C-Means:* memberikan nilai derajat keanggotaan pada setiap data terhadap klaster,

menghitung pusat klaster baru, dan melakukan pembaruan hingga kondisi berhenti terpenuhi.

6. Evaluasi Hasil Klasterisasi
   Validasi dilakukan menggunakan nilai *Silhouette Coefficient* untuk mengukur seberapa baik data tergabung dalam klaster masing-masing.

Tujuan utama penelitian ini adalah mengelompokkan Penyakit Menular dan Tidak Menular berdasarkan karakteristik tertentu sehingga hasilnya dapat digunakan sebagai bahan pertimbangan dalam pengambilan keputusan di Puskesmas Bayung Lencir. Data yang digunakan dalam penelitian ini merupakan data Penyakit TBC, DBD, Diare, dan Hipertensi yang didapat dari Puskesmas Bayung Lencir periode 2025. Data Penyakit yang digunakan untuk penelitian ini dapat dilihat pada Tabel 2.

Tabel 2. Data Penyakit Perwilayah

No	Wilayah	jumlah_kasus_penyakit_DBD	jumlah_kasus_penyakit_TBC	jumlah_kasus_penyakit_Hipertensi	jumlah_kasus_penyakit_Diare
1	MUARA MERANG	2	3	7	16
2	KALI BERAU	1	1	4	15
3	KEPAYANG	1	4	8	15
4	TAMPANG BARU	1	5	3	13
5	PAGAR DESA	1	1	2	10
6	LUBUK HARJO	1	3	9	41
7	SIMPANG BAYAT	8	9	3	5
8	MANGSANG	5	2	6	5
9	SINDANG MARGA	1	3	8	23
10	MENDIS JAYA	1	9	6	14
11	PANGKALAN BAYAT	5	1	5	21
12	MUARA BAHAR	1	7	3	17
13	MUARA MEDAK	4	1	5	22
14	MENDIS	6	1	3	13
15	TELANG	2	1	1	24
16	PULAI GADING	3	5	3	15
17	BAYUNG LENCIR	1	3	17	14
18	SENAWAR JAYA	1	4	5	23
19	SUNGAI NAPAL	1	2	3	16

Tahapan penelitian meliputi beberapa langkah. Pertama, pengumpulan data dari puskesmas. Kedua, melakukan pra-pemrosesan data berupa pembersihan data, normalisasi agar berada pada rentang tertentu, serta penanganan data hilang. Ketiga, menentukan parameter awal algoritma, seperti jumlah *cluster*, parameter pembobot m=2, nilai *error threshold*, dan iterasi maksimum. Keempat, menerapkan algoritma *K-Means* dan *Fuzzy C-Means* hingga tercapai konvergensi. Hasil klasterisasi diinterpretasikan ke dalam dua kelompok risiko, yaitu rendah dan tinggi. Seluruh proses penelitian ini dilakukan dengan bantuan *tools Jupyter Notebook* dan bahasa pemrograman *Python*. Digunakan pula perhitungan manual pada *Microsoft Excel* sebagai perbandingan hasil.

3. Coding dan Hasil Pembahasan

Berikut adalah *coding* pada *Python*:

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
from fcmeans import FCM
from sklearn.preprocessing import MinMaxScaler
```

```
    from sklearn.metrics import silhouette_score
# 1. Load dataset
    data = pd.read_excel("data_penyakit.xlsx")
    X = data[['DBD','TBC','Hipertensi','Diare']].values
# 2. Normalisasi data
    scaler = MinMaxScaler()
    X_scaled = scaler.fit_transform(X)
# 3. K-Means Clustering
    kmeans = KMeans(n_clusters=2, random_state=0).fit(X_scaled)
    data['Cluster_KMeans'] = kmeans.labels_
```

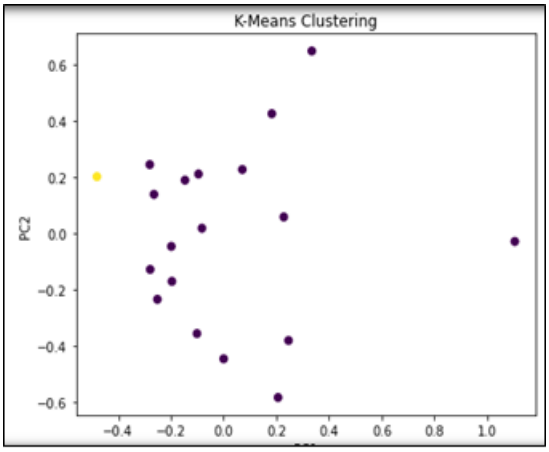
Dengan menerapkan metode *K-Means* dan *Fuzzy C-Means* pada pemrograman *Python dan tools Jupyter Notebook* untuk mengelompokkan data penyakit menular dan tidak menular dan menggunakan bantuan perhitungan manual pada *Microsoft Excel*, data yang berjumlah 19 dibagi menjadi 2 *Cluster*, yaitu *Cluster* Risiko Penyakit Cukup Rendah, dan *Cluster* Risiko Penyakit Cukup Tinggi. Hasil klasterisasi dapat dilihat pada Tabel 3.

Tabel 3. Analisis Hasil

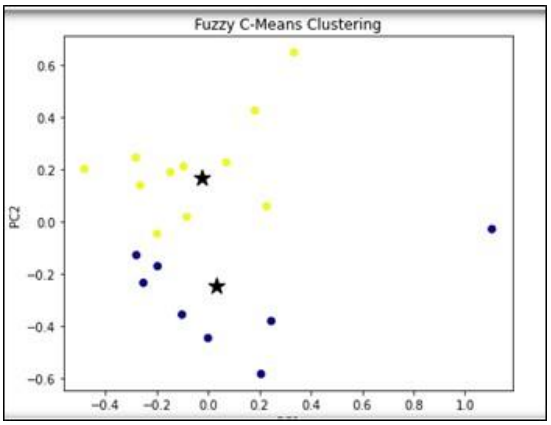
	K-Means		*Fuzzy C-Means*	
Hasil	C1	C2	C1	C2
	1 Data	18 Data	8 Data	11 Data
Iterasi	Manual	*Python*	Manual	*Python*
	2 Iterasi	2 Iterasi	27 Iterasi	26 Iterasi

Pada program *Python* dan perhitungan manual *K-Means,* perhitungan berhenti di iterasi yang sama, yaitu iterasi ke 2. Sedangkan pada *Fuzzy C-Means*, perhitungan manual berhenti pada iterasi 27, sedangkan pada program *python* berhenti pada itersi 26. Data-data digolongkan ke dalam masing-masing *cluster*, baik itu Risiko Penyakit Cukup Rendah, dan Risiko Penyakit Cukup Tinggi.

Berikut adalah tampilan pada Phyton untuk dapat melihat penyebaran data, maka dibutuhkan visualisasi grafik berupa *scatter plot* pada metode *K-Means dan Fuzzy C-Means*. Dapat dilihat pada Gambar 3 dan 4.



Gambar 3. Visualisasi *scatter plot K-Means*



Gambar 4. Visualisasi *scatter plot Fuzzy C-Mean*

Pada grafik visualisasi di atas, dapat dilihat bahwa penyebaran datanya cukup baik dengan bintang sebagai penanda *Centroid* atau Pusat *Cluster*-nya. Dengan visualisasi ini juga dapat dikatakan bahwa metode *Fuzzy C-Means* dapat bekerja dengan baik dalam pengelompokkan data.

Berdasarkan hasil pengelompokan menggunakan metode *K- Means*, diperoleh dua cluster utama yang merepresentasikan perbedaan distribusi persentase kasus penyakit pada wilayah penelitian. Cluster 0, yang terdiri dari 18 wilayah, memperlihatkan persentase kasus penyakit yang sangat tinggi pada keempat kategori penyakit yang dianalisis, yaitu Demam Berdarah Dengue (DBD), Tuberkulosis (TBC), Hipertensi, dan Diare. Persentase tertinggi ditunjukkan oleh kasus DBD sebesar 97,83%, diikuti oleh TBC sebesar 95,38%, Hipertensi sebesar 91,09%, dan Diare sebesar 87,27%.

Sementara itu hasil pengelompokan menggunakan metode *Fuzzy C-Means* menghasilkan dua cluster dengan karakteristik distribusi kasus penyakit yang berbeda. Cluster 0, yang terdiri dari 8 wilayah, menunjukkan persentase kasus penyakit yang relatif lebih tinggi pada DBD sebesar 69,57%, sedangkan penyakit lainnya seperti TBC (26,15%), Hipertensi (28,71%), dan Diare (35,71%) berada pada tingkat yang lebih rendah.

4. Kesimpulan

Penelitian ini bertujuan untuk menganalisis dan mengelompokkan penyakit menular dan tidak menular berdasarkan wilayah tempat tinggal di Puskesmas Bayung Lencir menggunakan metode *K-Means* dan *Fuzzy C-Means (FCM)*. Berdasarkan hasil pengujian, kedua metode mampu membentuk dua klaster risiko penyakit, yaitu risiko rendah dan risiko tinggi, sesuai dengan karakteristik data penyakit DBD, TBC, Hipertensi, dan Diare.

Proses perhitungan secara manual dan pada program *Python* menunjukkan hasil *clustering* yang konsisten pada pengelompokan akhir pada masing-masing metode, walaupun terdapat perbedaan pada iterasi di masing-masing metode: Manual *Fuzzy C-Means* berhenti di iterasi ke-27, *python* berhenti di iterasi ke-26. Manual *K-Means* berhenti di iterasi ke-2, *python* berhenti di iterasi ke-2.

Klasterisasi membagi 19 data penyakit menular dan tidak menular baik itu pada perhitungan manual *Fuzzy C-Means* ataupun *Python* menjadi: *Cluster* risiko penyakit cukup rendah: 8 data, *Cluster* risiko penyakit cukup tinggi: 11 data. pada perhitungan manual *K-Means* ataupun *Python* menjadi: *Cluster* risiko penyakit cukup rendah: 1 data, *Cluster* risiko penyakit cukup tinggi: 18 data.

Sistem analisis yang dibangun mampu membantu pengelompokan data penyakit menular maupun penyakit tidak menular dengan akurasi yang cukup baik, sehingga berpotensi menjadi alat bantu pengambilan keputusan dalam bidang kesehatan masyarakat.

Hasil clustering menunjukkan adanya distribusi wilayah yang berbeda untuk masing-masing jenis penyakit, sehingga dapat digunakan sebagai dasar dalam menentukan prioritas penanganan kesehatan di tiap wilayah.

Penelitian selanjutnya disarankan untuk menambah jumlah data dan variabel lingkung penduduk, sumber air bersih, serta perilaku hidup bersih dan sehat (PHBS), agar model

klasterisasi dapat memberikan hasil yang lebih akurat dan komprehensif.

Daftar Pustaka

[1] H. Mubarak dan G. Kholijah, "Analisis Klaster Dalam Pengelompokan Kabupaten/Kota Di Provinsi Jambi Berdasarkan Penyakit Menular Menggunakan Metode K-Means," *J. Statistika dan Komputasi*, vol. 2, no. 1, pp. 1–20, Jun. 2023.

[2] A. Rusmiati, *Skripsi Optimasi K-Means Clustering pada Penyakit Menular di Puskesmas Bontomarannu*, Univ. Hasanuddin, 2023.

[3] A. Kurnia, "Perbandingan Algoritma K-Means dan Fuzzy C-Means Untuk Clustering Puskesmas Berdasarkan Gizi Balita Surabaya," *Processor*, vol. 18, no. 1, Apr. 2023.

[4] R. Chaniago, *Pengelompokan Daerah Penyebaran Penyakit Menular di Kota Sibolga Menggunakan Algoritma Fuzzy C-Means*, Univ. Malikussaleh, 2024.

[5] M. H. Ziqrullah et al., "Clustering Data Penyakit Pasien Pada Puskesmas Mulyaguna Menggunakan Algoritma K-Means," *JUMIN*, vol. 6, no. 2, pp. 608–616, Des. 2024.

[6] D. Puspitasari, "Analisis Perbandingan Algoritma K-Means dan Fuzzy C-Means dalam Pengelompokan Data Penyakit Menular," Jurnal Teknologi dan Sistem Informasi, vol. 9, no. 2, pp. 85–92, 2022.

[7] M. N. Fadilah, A. W. Hidayat, and L. Marlina, "Implementasi Algoritma Fuzzy C-Means untuk Pengelompokan Daerah Rawan Penyakit Demam Berdarah," Jurnal Informatika dan Sistem Cerdas (JISC), vol. 5, no. 2, pp. 56–64, 2022.

[8] Dinas Kesehatan Kabupaten Musi Banyuasin, Laporan Tahunan Kasus Penyakit Menular di Wilayah Puskesmas Bayung Lencir Tahun 2023, Sekayu: Dinkes Muba, 2023.

[9] A. Lesmana, "Penerapan Logika Fuzzy K-Means untuk Clustering Data Obat pada Puskesmas Pematang Johar," UMA, 2023.

[10] Skripsi, "Penerapan Analisis Cluster dengan K-Means dalam Mengelompokkan Jumlah Kematian COVID-19 di Indonesia," UNM, 2023.

[11] W. Aulia, "Analisis Algoritma K-Means Clustering pada Penyakit Tidak Menular," *JSSR*, 2025.

[12] Y. Febriani, "Metode K-Means Cluster Untuk Mengelompokkan Kota," *Sainmatika*, 2021.

[13] Marince Cristina Panjaitan, "Analisa Perbandingan Algoritma Clustering untuk Pemetaan Status Gizi Balita di Puskesmas Pasir Jaya," *ResearchGate*, 2025.

[14] A. Rusmiati, *Optimasi K-Means Clustering pada Penyakit Menular di Puskesmas Bontomarannu*, Univ. Hasanuddin, 2023.

[15] A. Kurnia, "Perbandingan Algoritma K-Means dan Fuzzy C-Means untuk Clustering Puskesmas Berdasarkan Gizi Balita," *Processor*, vol. 18, no. 1, Apr. 2023.

[16] R. Chaniago, *Pengelompokan Daerah Penyebaran Penyakit Menular di Kota Sibolga Menggunakan Algoritma Fuzzy C-Means*, Univ. Malikussaleh, 2024.

[17] N. D. A. L. R. Putri, "Analisa Perbandingan K-Means dan Fuzzy C-Means dalam Pengelompokan Daerah Penyebaran COVID-19 Indonesia," *Sistemasi*, vol. 10, no. 2, 2021.

[18] G. A. Budianto, "Evaluasi K-Means dan Fuzzy C-Means Clustering: Analisis Data HIV di Kota Surabaya," *Jurnal Kesehatan Informasi*, 2023.

[19] M. H. Sholeh et al., "Implementasi K-Means Clustering dalam Pemetaan Wilayah Rawan Penyakit Leptospirosis di Jawa Timur," *JRAM*, 2024.

[20] S. F. Octavia dan M. Mustakim, "Penerapan K-Means dan Fuzzy C-Means untuk Pengelompokan Data Kasus COVID-19 di Kabupaten Indragiri Hilir," *BITS*, vol. 3, no. 2, 2021.