

Pengelompokan Kasus *Tuberculosis* Dengan Algoritma *K-Means* Berdasarkan Kelurahan di Kota Bogor

Isti Juliana Putri^{1*}, Freza Riana², Berlina Wulandari³

^{1,2,3} Universitas Ibn Khaldun Bogor
Jl. Sholeh Iskandar Kedungbadak Tanah Sereal Kota Bogor, Indonesia

E-mail korespondensi : istijuliputri34@gmail.com

Informasi Artikel	Diterima: 11-11-2023	Direvisi: 11-03-2024	Disetujui: 03-04-2024
-------------------	----------------------	----------------------	-----------------------

Abstrak

Tuberculosis (TBC) merupakan penyakit menular yang masih menjadi masalah sampai saat ini, angka kesakitan dan kematian akibat bakteri ini pun sangat tinggi. Penyakit TBC masih menjadi masalah kesehatan baik di dunia maupun di Indonesia terkhususnya di Jawa Barat Kota Bogor. Dengan luas wilayah Kota Bogor sebesar 11.850 Ha, tidak bisa dipungkiri bahwa terdapat penyebaran penyakit menular di Kota Bogor, penyakit menular di Kota Bogor masih cukup tinggi yang didominasi oleh penyakit TBC, DBD, dan HIV. Dengan banyaknya data kasus penyakit TBC di Kota Bogor diperlukan pengelompokan penyakit TBC untuk mengetahui wilayah kelurahan mana saja yang kasus TBCnya itu tinggi atau rendah. Diperlukan metode yang cepat dan akurat untuk menentukan *cluster* di suatu daerah, karena dapat menjadi salah satu kunci pencegahan atau penyuluhan terkait penyakit TBC. Salah satu metode *clustering* adalah algoritma *K-Means* algoritma ini dapat membagi data menjadi satu atau lebih *cluster* dengan karakteristik yang mirip, metode ini mampu mencapai akurasi serta kecepatan prosesnya juga relatif tinggi. Penelitian ini diharapkan dapat digunakan bagi Dinas Kesehatan Kota Bogor sebagai data rujukan untuk menindak lanjuti penyebaran penyakit TBC di Kota Bogor. Berdasarkan hasil penelitian menggunakan algoritma *K-Means* diperoleh evaluasi dengan *Silhouette Coefficient* dipilih dua *cluster* karena memiliki nilai yang paling tinggi. Kelompok penyakit TBC tertinggi berada di *cluster* 2 dengan jumlah 22 kelurahan di Kota Bogor.

Kata Kunci: Algoritma K-Means, Kota Bogor, Tuberculosis.

Abstract

Tuberculosis (TB) is an infectious disease that is still a problem today, the morbidity and mortality rate due to this bacteria is very high. TB is still a health problem both in the world and in Indonesia, especially in West Java, Bogor City. With an area of 11,850 hectares, it is undeniable that there is a spread of infectious diseases in Bogor City, infectious diseases in Bogor City are still quite high, dominated by tuberculosis, dengue fever, and HIV. With so much data on TB cases in Bogor City, it is necessary to group TB diseases to find out which urban villages have high or low TB cases. A fast and accurate method is needed to determine clusters in an area, because it can be one of the keys to prevention or counseling related to TB disease. One of the clustering methods is the K-Means algorithm, this algorithm can divide data into one or more clusters with similar characteristics, this method is able to achieve accuracy and the speed of the process is also relatively high. This research is expected to be used for the Bogor City Health Office as reference data to follow up on the spread of TB disease in Bogor City. Based on the results of research using the K-Means algorithm, evaluation with Silhouette Coefficient, two clusters were selected because they had the highest value. The highest TB disease group is in cluster 2 with a total of 22 villages in Kot.

Keywords: K-Means Algorithm, Bogor City, Tuberculosis.

1. Pendahuluan

Tuberculosis (TBC) merupakan penyakit menular yang disebabkan oleh kuman *Mycobacterium tuberculosis* yang menjangkit bagian organ paru-paru (Toresa, 2020) (Lenie M et al., 2019). Pengidap TBC mengalami berbagai gejala termasuk demam, kurus, dan batuk

(Sulistiyono Hidayat et al., 2022). TBC merupakan masalah kesehatan baik di dunia maupun di Indonesia yang kasus dan kematiannya tinggi oleh karena itu TBC di Indonesia termasuk program unggulan pemerintah untuk diatasi. (Dinas Kesehatan Kota Bogor., 2022) (Jenita DT, D et al., 2020). Di Indonesia, TBC paru merupakan



penyakit yang ditakuti karena membunuh baik anak-anak maupun orang dewasa. Secara umum penyakit TBC paru disebabkan oleh beberapa faktor yaitu lingkungan, cuaca, tempat tinggal dan keturunan. Penyebaran penyakit TBC paru khususnya tersebar luas di Indonesia termasuk di Jawa Barat di Kota Bogor (Febriyanti et al., 2020).

Luas wilayah Kota Bogor sebesar 11.850 Ha yang terdiri dari 6 kecamatan dan 68 kelurahan (Diskominfo Kota Bogor., 2021), dengan wilayah yang cukup luas tidak bisa dipungkiri bahwa terdapat penyebaran penyakit menular di Kota Bogor, penyakit menular di Kota Bogor masih cukup tinggi yang didominasi oleh penyakit TBC, DBD, dan HIV (Diskominfo Kota Bogor., 2019). Dengan banyaknya data kasus penyakit TBC di Kota Bogor, pengelompokan penyakit TBC menurut sub wilayah diperlukan untuk mengetahui sub wilayah mana saja yang paling banyak kasus penyakit TBC. Diperlukan metode yang cepat dan akurat untuk menentukan *cluster* di suatu daerah, karena dapat menjadi salah satu kunci pencegahan atau penyuluhan terkait penyakit TBC.

Metode *clustering* adalah salah satu dari sekian banyak metode yang ada di data mining. Metode ini efektif dalam menghasilkan *cluster* di tumpukan data, atau mengelompokkan data ke dalam beberapa *cluster*, sehingga data yang berbeda harus menjadi bagian dari *cluster* lain (Minarni et al., 2021) (Benabdellah et al., 2019). Kerja teknik *clustering* didasarkan pada fakta bahwa populasi atau bagian dari data yang kita miliki dikelompokkan menjadi beberapa *cluster* sehingga titik data dalam kelompok yang sama lebih mirip dengan titik data lain dalam kelompok yang sama dibandingkan dengan kelompok lain (Sulistiyo Hidayat et al., 2022) (Wardani et al., 2019).

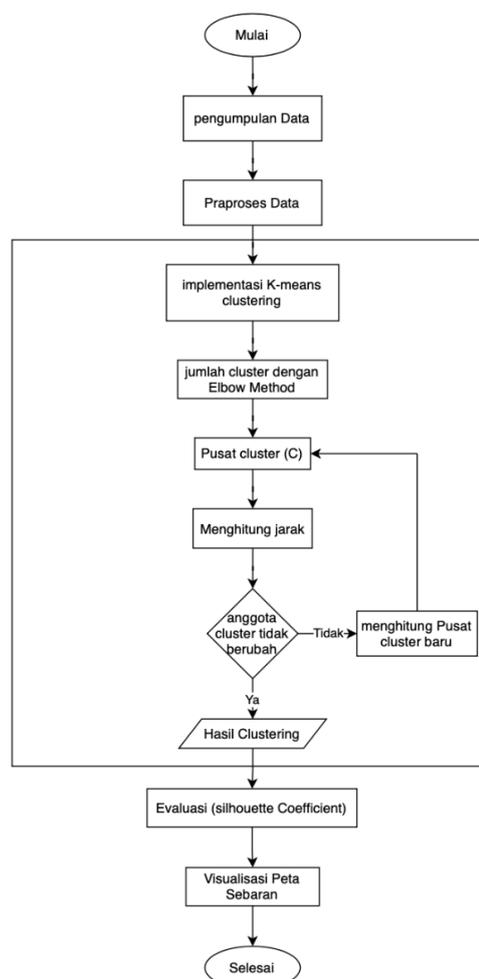
Pada penelitian terkait Implementasi Algoritma *K-Means* Dalam Pengelompokan Kasus Penyakit Tuberculosis Paru Berdasarkan Provinsi, data yang digunakan yaitu data penyakit TBC paru pada tahun 2019, dengan menghasilkan 3 kategori *cluster* tinggi, sedang, dan rendah. *cluster* tinggi berisi 3 provinsi, *cluster* sedang berisi 3 provinsi, dan *cluster* rendah berisi 28 provinsi. Pada penelitian ini terbukti bahwa Algoritma *K-Means* dapat membagi data menjadi satu atau lebih *cluster* dengan karakteristik yang sama dengan jumlah *cluster* yang sudah ditetapkan. Algoritma ini banyak digunakan karena sederhana, prosesnya dan konvergensinya yang cepat (Febriyanti et al., 2020) (Chunhui Y, Haitao Y., 2019)

Maka dari itu untuk mengetahui pengelompokan penyakit TBC menurut sub wilayah kelurahan di Kota Bogor penelitian ini menggunakan algoritma *K-Means*, sehingga penelitian ini bertujuan untuk menghasilkan

wilayah mana saja yang kasus TBC dari level tertinggi hingga rendah, selanjutnya dilakukan evaluasi dengan *silhouette coefficient* agar dapat menentukan jumlah *cluster* terbaik yang akan diambil. Selanjutnya hasil akan divisualisasikan dalam bentuk peta sebaran penyakit TBC di Kota Bogor. Hasil penelitian ini dimaksudkan sebagai rekomendasi maupun upaya bagi DINKES Kota Bogor untuk menerapkan pencegahan dan penyuluhan terkait penyakit TBC di daerah-daerah yang ada di Kota Bogor, Pemerintah nantinya dapat lebih fokus pada daerah dengan kasus TBC yang tinggi, dan bersiaga untuk sub wilayah kelurahan yang jumlah warganya banyak agar penyebaran penyakit TBC tidak cepat menyebar.

2. Metode Penelitian

Metode penelitian ini dapat dilihat pada gambar berikut.



Gambar 1. Metode Penelitian

Metode pada penelitian ini dimulai dari:

2.1. Pengumpulan data

Menggunakan data pasien TBC dari Dinas Kesehatan (DINKES) Kota Bogor pada

tahun 2021, dan data penduduk diperoleh secara online pada situs web resmi Badan Pusat Statistik Kota Bogor (BPS) pada alamat: <https://bogorkota.bps.go.id/>.

2.2. Praproses data

Adapun tahapan praproses data yaitu, data *selection* yaitu memilih data atau variabel yang akan digunakan, data *cleaning* yaitu pembersihan data yang kosong atau duplikat, rasio data yaitu mencari nilai rasio yang dihitung dengan rumus:

$$\frac{\text{Rasio Teridentifikasi}}{\text{jumlah warga}} = \frac{\text{jumlah teridentifikasi}}{\text{jumlah teridentifikasi}} \quad (1)$$

$$\frac{\text{Rasio Positif TBC}}{\text{jumlah warga}} = \frac{\text{jumlah positif TBC}}{\text{jumlah positif TBC}} \quad (2)$$

Dan normalisasi data untuk meminimalisi data null atau menskalakan data yang sama (Marlina et al., 2018).

$$v' = \frac{v - \min_A}{\max_A - \min_A} \quad (3)$$

Keterangan:

v' = nilai normalisasi data
 v = data asli
 \min_A = nilai minimum variabel A
 \max_A = nilai maksimum variabel A

2.3. Implementasi K-Means clustering.

Clustering adalah studi formal tentang metode dan algoritma untuk mengelompokkan atau klusterisasi (Mehmed Kantardzic, 2020a).

Algoritma pengelompokan partisi *K-Means* yang sederhana sangat efisien secara komputasi dan memberikan hasil yang sangat bagus jika *clusternya* kompak, berbentuk *hyperspherical*, dan terpisah dengan baik dalam ruang fitur. Adapun Langkah-langkah melakukan *clustering* dengan metode *K-Means* adalah sebagai berikut:

1. Menentukan jumlah *cluster* yang ingin dibentuk. Untuk menentukan jumlah cluster pada penelitian ini menggunakan metode *elbow*.

2.4. Metode Elbow

Metode *elbow* merupakan metode optimasi kluster yang membentuk siku pada suatu titik, yaitu titik di mana penurunan varians intra-kluster menjadi lebih lambat setelahnya. Untuk mengetahui perbandingannya adalah dengan menghitung nilai *Sum Of Square Error* (SSE) dari masing-masing kluster (Ayu et al., 2019).

$$SSE = \sum_{\text{cluster}=1}^{\text{cluster}} \sum_{xi} |xi - C|^2 \quad (4)$$

Keterangan:

C = pusat *cluster* ke- c
 X_i = jarak data obyek ke- i
 C = pusat *cluster* ke- i

2. Memilih nilai untuk pusat *centroid* (C) awal, untuk menentukan nilai *centroid* dilakukan secara acak.
 3. Menghitung jarak setiap data input terhadap masing-masing *centroid* menggunakan rumus jarak *Euclidian Distance*:

$$d(ai, bi) = \sqrt{\sum_{i=1}^n (ai - bi)^2} \quad (5)$$

Keterangan:

$d(ai, bi)$: jarak objek antara nilai data dan *centroid*
 ai : nilai data dari dimensi
 bi : nilai *centroid* dari dimensi
 n : banyaknya dimensi atau atribut data

4. Mengelompokkan setiap data berdasarkan kedekatannya dengan *centroid* (jarak terkecil).
 5. Memperbarui nilai *centroid* (C). Apabila anggota dari setiap *cluster* berubah dari perhitungan iterasi pertama maka nilai *centroidnya* diperbarui dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus:

$$C = \frac{\sum m}{N} \quad (6)$$

Keterangan:

C : *centroid* data
 m : jumlah nilai jarak yang masuk dalam setiap *cluster*
 N : jumlah data dalam *cluster*
 6. Ulangi Langkah ke 3 s/d 5 hingga nilai dari setiap anggota tidak berubah (Mehmed Kantardzic, 2020b).

2.5. Evaluasi dengan Silhouette coefficient

Pada tahapan *silhouette coefficient* ini digunakan di akhir untuk menyediakan informasi tentang kualitas dan kekuatan pada setiap hasil *cluster* pada proses *clustering*, seberapa optimal atau tidaknya suatu obyek ditempatkan dalam suatu *cluster*.

$$SI = \frac{1}{\text{cluster}} \sum_{j=1}^{\text{cluster}} SI_j \quad (7)$$

Keterangan:

SI = rata-rata *Sillhouette* index dari data
 SI_j = rata-rata *Sillhouette* index *cluster* j
 $Cluster$ = jumlah *cluster*

(Ayu et al., 2019) (Solmin P, Hisma A. 2021).

3. Hasil dan Pembahasan

Hasil dari analisis dalam penelitian ini dijabarkan sebagai berikut:

3.1 Pengumpulan Data

Data yang digunakan pada penelitian ini yaitu, data pasien TBC terkumpul sebanyak 2916 dan data penduduk dari 68 kelurahan sebanyak 1.091396 di Kota Bogor pada tahun 2021. Adapun data awal penyakit TBC terdapat 6 variabel sebagai berikut:

1. Tanggal daftar pasien.
2. Bulan daftar pasien.
3. Fasyankes, yang terbagi 2 sub yaitu jenis fasyankes dan nama fasyankes.
4. Alamat terduga TBC, yang terbagi menjadi 3 sub yaitu, kabupaten/kota, kecamatan dan kelurahan.
5. Anatomi penyakit.
6. Hasil diagnosis.

Tabel 1. Data penduduk sumber BPS Kota Bogor 2021

No	Nama kelurahan	Jumlah warga tahun 2021
1	Balungbangjaya	13077
2	Bubulak	17828
3	Cilendek barat	20266
...

Sehingga data yang digunakan adalah sebagai berikut:

Tabel 2. Data yang digunakan

Kelurahan	Jumlah warga	Anatomi	Hasil diagnosis
Balungbangjaya	13077	67	44
Bubulak	17828	120	69
Cilendek barat	20266	62	49
Cilendek timur	18924	46	40
...

3.2. Praproses Data

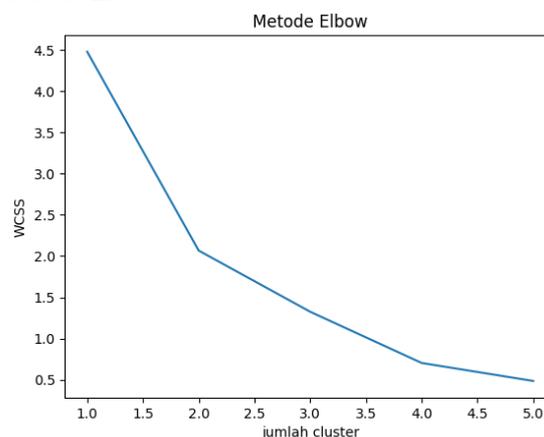
Setelah dilakukan tahapan praproses data, dihasilkanlah data yang siap digunakan untuk analisis dengan algoritma k-means, terdapat perubahan penamaan pada variabel tertentu seperti anatomi berubah menjadi teridentifikasi hasil diagnosis menjadi positif TBC. Adapun data hasil praproses yaitu sebagai berikut:

Tabel 3. Hasil Praproses data

Kelurahan	Normalisasi rasio teridentifikasi	Normalisasi rasio positif
Balungbangjaya	0,669105431	0,342820499
Bubulak	0,897559104	0,400617948
Cilendek barat	0,375744527	0,234588166
...

3.3 Implementasi K-Means Clustering

Tahap pertama dalam analisis dengan *k-means*, yaitu menentukan jumlah cluster yang akan dipilih. Pada tahap ini dilakukan dengan metode elbow dan dihasilkan cluster yang baik berada di area patahan yang paling cekung yaitu cluster 2.



Gambar 2. Elbow Method

Selanjutnya yaitu tahap kedua menentukan pusat cluster (centroid). Dalam penelitian ini dipilih centroid secara acak dan didapatkan centroid sebagai berikut:

Tabel 4. Nilai Centroids

Kelurahan	Teridentifikasi	Positif TBC
Baranangsiang	0,462783972	0,340145697
Cikaret	0,487543968	0,347750594

Jadi untuk nilai *centroid* (C) awal ini yaitu:

C1= 0,462783972, 0,340145697

C2= 0,487543968, 0,347750594

Tahap ketiga menghitung dengan Euclidian distance dan termasuk kedalam cluster satu atau dua, untuk menentukan termasuk ke dalam cluster berapa yaitu dilihat dari nilai terkecil dari tiap kelurahannya, hasilnya sebagai berikut:

Tabel 5. Hasil cluster

Kelurahan	C1	C2	Cluster
Balungba	0,391424	0,01292	Cluster
ngjaya	452	4606	2
Bubulak	0,626747	0,22523	Cluster
	501	9232	2
...
Sukaresm	0,097196	0,30937	Cluster
i	033	3373	2
Tanahsar	0,070683	0,39435	Cluster
eal	152	5819	2

Dan dihasilkan pada iterasi ke-1:

Cluster 1= 46 kelurahan

Cluster 2= 22 kelurahan

Tahap selanjutnya dilakukan seperti diatas secara berulang sampai nilai centroid dan keanggotaan dari setiap cluster tidak berubah. Pada penelitian ini berhenti di iterasi ke-2 dengan memperbarui nilai centroids, hasil sebagai berikut:

Tabel 6. Nilai centroid baru

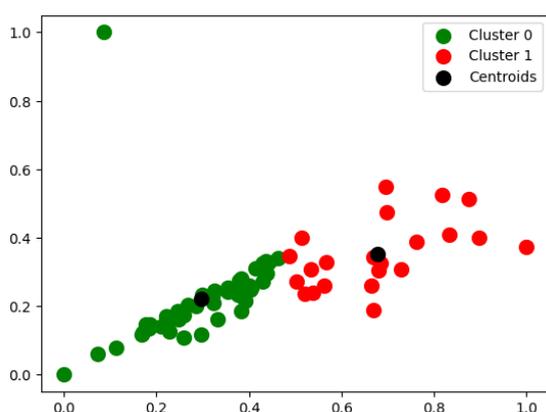
Centroids ke-2	Teridentifikasi	Positif TBC
C1	0,296894975	0,221690495
C2	0,677484452	0,352661101

Hasil dari iterasi ke-2:

Cluster 1 = 46 kelurahan

Cluster 2 = 22 kelurahan

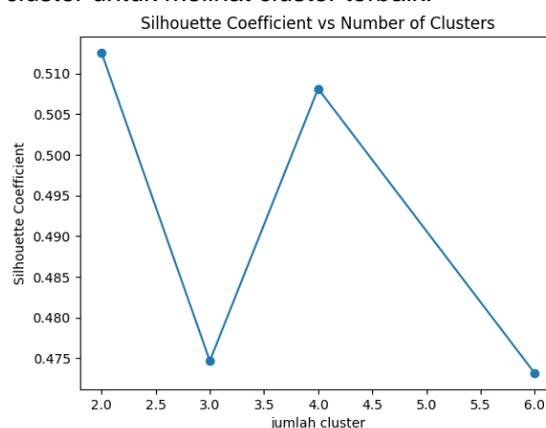
Karena hasil dari setiap *cluster* tidak berubah dan nilai centroidnya tidak berubah maka pada penelitian ini berhenti pada iterasi ke-2. Adapun hasil perhitungan k-means dengan bantuan python ini dapat divisualisasikan dengan scatter plot, dimana sebaran kelurahan berdasarkan jumlah kasus TBC yang dihasilkan.



Gambar 3. Hasil Scatter Plot

3.4 Evaluasi Silhouette Coefficient

Pada tahap evaluasi menggunakan silhouette coefficient ini diuji cluster sebanyak 6 cluster untuk melihat cluster terbaik.



Gambar 4. Hasil Silhouette Coefficient

Hasil dari evaluasi dengan silhouette coefficient ini dipilih dari nilai tertinggi dan didapatkan nilai tertinggi berada di cluster 2 dengan hasil 0.5125 maka dalam penelitian ini dipilih cluster terbaik berada di 2 cluster. Maka Dengan keanggotaan dari 2 cluster sebagai berikut:

Tabel 7. Hasil dari *cluster* 1

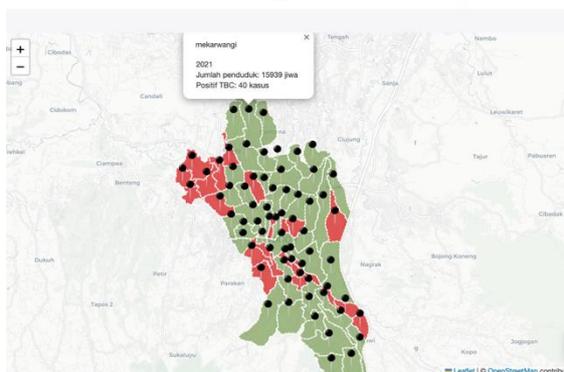
Jenis <i>cluster</i>	Nama Kelurahan	Kecamatan
Cluster 1	Cilendek barat, Cilendek timur, Curugmekar, Gunung batu, Loji, Menteng, Pasirmulya	Bogor barat
	Bojong kerta, Cipaku, Empang, Genteng, Harjasari, Kertamaya, Lawanggintung, Muarasari, Mulyaharja, Pakuan, Pamoyanan, Rancamaya, Ranggamekar	Bogor Selatan
	Babakan pasar, Ciwaringin, Kebon kelapa, Pabaton, Paledang, Panaragan, Tegallega	Bogor Tengah
	Baranangsiang, Katulampa	Bogor Timur
	Bantarjati, Cibuluh, Ciluar, Ciparigi, Kedunghalang, Tanah baru, TegalGundil	Bogor Utara
	Cibadak, Kayumanis, Kebon pedes, Kedungbadak, Kedungjaya, Kencana, Mekar, wangi, Sukadamai, Sukaresmi, Tanahsareal	Tanah sareal

Tabel 8. Hasil dari cluster 2

Jenis Cluster	Nama kelurahan	Kecamatan
Cluster 2	Balungbangjaya, Bubulak, Curug, Margajaya, Pasirjaya, Pasirkuda, Semplak, Sindangbarang, Situ gede	Bogor Barat
	Batutulis, Bondongan, Cikaret	Bogor Selatan
	Babakan, Cibogor, Gudang, Sempur	Bogor Tengah
	Sindangrasa, Sindangsari, Sukasari, Tajur	Bogor Timur
	Cimahpar	Bogor Utara
	Kedung waringin	Tanah sereal

3.5 Visualisasi Peta Sebaran

Visualisasi Peta Sebaran Penyakit TBC di Kota Bogor



Gambar 5 Hasil visualisasi peta sebaran

Hasil visualisasi dari peta sebaran ini dikategorikan, wilayah berwarna hijau kategori rendah dengan jumlah 46 kelurahan, dan berwarna merah kategori tinggi berjumlah 22 kelurahan. Hasil dari kategori ini dilihat dari banyaknya jumlah kasus TBC dari setiap kelurahan di Kota Bogor pada tahun 2021.

4. Kesimpulan

Pada penelitian ini dihasilkan pengelompokan menurut subwilayah, menurut hasil dari evaluasi *silhouette coefficient* yaitu dipilih *cluster* terbaik berada di *cluster 2*, dimana: *Cluster 1*, termasuk ke dalam kategori *cluster* rendah yang berjumlah 46 kelurahan, dengan karakteristik jumlah teridentifikasi yang rata-rata rendah, jumlah positif TBC yang rata-rata rendah. Tetapi ada satu kelurahan yang Pencilan karena memiliki nilai paling tinggi, dipengaruhi oleh jumlah warga yang sangat banyak atau kelurahan tersebut padat penduduk dikarenakan kelurahan tersebut berada di pusat Kota Bogor, yaitu kelurahan Paledang. *Cluster 2*, termasuk ke dalam kategori tinggi yang berjumlah 22 kelurahan, dengan karakteristik jumlah teridentifikasi yang rata-rata tinggi dan jumlah positif TBC yang rata-rata tinggi. Akan tetapi ada kelurahan yang berkarakteristik jumlah warga yang rendah, hal itu dipengaruhi jumlah teridentifikasi dan positif TBC yang tinggi.

Dari hasil analisis pengelompokan ini diambil berdasarkan banyaknya kasus TBC di Kota Bogor yang dipengaruhi oleh banyaknya jumlah teridentifikasi dan jumlah penduduk di kota bogor pada tahun 2021. Dari hasil penelitian ini juga menunjukkan bahwa *clustering* dengan algoritma *K-Means* terbukti dapat mengelompokkan beberapa data menjadi ke dalam satu kelompok berbeda yang terdiri dari beberapa data acak, sesuai dengan kedekatan jarak centroidnya.

Referensi

- Toresa, D (2020). Implementasi K-Means Terhadap penyebaran penyakit Tbc di Riau menggunakan Rapid Miner. *JUTIM (Jurnal Teknik Informatika Musirawas)*, 5(1), 35-34.
- Lenie M, S.KM. M.KL., Dr. dr. H. Syamsul A, M.PD, DLP., Ihya Hazairin N, S.KM, M.PH., Atikah R, S.KM, M.PH., Dr. Tien Z, S.KM, M.KL., Agung W, S.T, M.T. (2019). *Desain Kemandirian Pola Perilaku Kepatuhan Minum Obat Pada Penderita TB Anak Berbasis Android*. CV Mine, 2-4.
- Dinas Kesehatan Kota Bogor. (2022). *Untuk Meningkatkan Notifikasi Kasus TBC di Kota Bogor, Perlunya Peningkatan Jejaring RS di Kota Bogor*. Dinas Kesehatan Kota Bogor.
- Jenita DT, D., Eva Y, Bondan P, Bambang T, Rosa Delima E, Lucky H. (2020). *Menuju Desa Sehat Bebas Tuberkulosis Pencegahan Tuberkulosis Melalui Program Pengembangan Desa Sehat (PPDS)*. Husada Mandiri, 1-2.
- Sulistiyo Hidayat, F., Berliana, R., Affandi, P., Zuliana, V., & Padilah, T. N. (2022). Penerapan K-Means Clustering dalam Pengelompokan Kasus Tuberkulosis di Provinsi Jawa Barat. *Jurnal Ilmiah Wahana Pendidikan*, 8(15), 213–227. <https://doi.org/10.5281/zenodo.7049113>
- Febriyanti, V., Satria Tambunan, H., Saragih, I. S., Damanik, I. S., Okprana, H., Tunas Bangsa, S., Jenderal, J., Blok, S., No, A., & Pematangsiantar, I. (2020). Implementasi Algoritma K-Means Dalam Pengelompokan Kasus Penyakit Tuberkulosis Paru Berdasarkan Provinsi. *Prosiding Seminar*

- Nasional Riset Dan Information Science (SENARIS)*, 2, 450–456.
- Diskominfo Kota Bogor. (2019). *Tren Penyakit Masih Tinggi, Dinkes Lakukan Tiga Program Ini*. Kota Bogor.Go.Id.
- Diskominfo Kota Bogor. (2021). *Letak Geografis*. Kota Bogor.Go.Id.
- Minarni, M., Sari, E. I., Syahrani, A., & Mandarani, P. (2021). Klasterisasi Penyakit Menggunakan Algoritma K-Medoids pada Dinas Kesehatan Kabupaten Agam. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 10(3), 137.
<https://doi.org/10.23887/janapati.v10i3.34904>
- Benabdellah, A. C., Benghabrit, A., & Bouhaddou, I. (2019). A survey of clustering algorithms for an industrial context. *Procedia Computer Science*, 148, 291–302.
<https://doi.org/10.1016/j.procs.2019.01.022>.
- Chunhui Yuan, Haitao Yang. (2019). Research on K-Value Selection Method of K-Means Clustering Algorithm. *J*, 2(2), 226-235.
- Wardani, R. S., Purwanto, Sayono, & Paramananda, A. (2019). Clustering tuberculosis in children using K-Means based on geographic information system. *AIP Conference Proceedings*, 2114.
<https://doi.org/10.1063/1.5112483>
- Mehmed Kantardzic. (2020). *Data Mining Concepts, Models, Methods, and Algorithms*. John Wiley, Sons, INC., Hoboken, New Jersey. Simultaneously in Canada.
- D. Ayu, I.C, Dewi, & K.Pramita. (2019). Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. *JURNAL MATRIX*, (9)3.
- Solmin Paembonan, Hisma Abduh. (2021). Penerapan Metode Silhouette Coeficient untuk Evaluasi Clustering Obat. *JURNAL ILMIAH ILMU-ILMU TEKNIK*, 2, (6), 50-51.
- Marlina, D., Lina, N., Fernando, A., & Ramadhan, A. (2018). Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi*, 4(2), 64.
<https://doi.org/10.24014/coreit.v4i2.4498>