

Penerapan Algoritma *K-Means Clustering* dan *Correlation Matrix* Untuk Menganalisis Risiko Penyebaran Demam Berdarah di Kota Pekanbaru

M. Azwan¹, Rahmad Kurniawan², Pizaini³, Fitri Insani⁴

^{1,2,3,4} Teknik Informatika, Universitas Islam Sultan Syarif Kasim Riau, Indonesia

¹11751100021@students.uin-suska.ac.id

²rahmadkurniawan@uin-suska.ac.id

³pizaini@uin-suska.ac.id

⁴fitri.insani@uin-suska.ac.id

Received : 24-11-2021; Accepted: 06-12-2021; Published: 25-12-2021

Abstrak— Kasus Demam Berdarah Dengue (DBD) di Kota Pekanbaru pada bulan November 2020 mencapai angka 2.788 kasus dan 33 kematian. Pemerintah telah melakukan sosialisasi pemberantasan sarang nyamuk dan menyediakan alat dan bahan pengendalian vektor. Namun, upaya pemerintah tersebut belum efektif karena metode yang diterapkan belum mampu mengacu kepada data dan informasi vektor. Machine learning dapat digunakan untuk menganalisis masalah khusus seperti DBD. Oleh karena itu penelitian ini menggunakan algoritma machine learning yaitu *k-means clustering* dan *correlation matrix* untuk analisis risiko DBD di Pekanbaru. Penelitian ini menggunakan data 12 kecamatan dan 50 atribut, serta data cuaca pada tahun 2020. *K-means* secara otomatis mencari cluster yang tidak diketahui dari data kasus DBD dengan cepat yang menghasilkan cluster C1 (Sukajadi, Senapelan), C2 (Tenayan raya, Tampan), dan C3 (Rumbai Pesisir, Rumbai). Berdasarkan pengujian eksperimental, penelitian ini menghasilkan nilai silhouette score 0,6. Sedangkan *Correlation matrix* mencari hubungan relevan yang tersembunyi didalam data. Hasil dari *correlation matrix* diperoleh hubungan linear yang kuat antara jumlah penduduk (JP) dan penderita (P) sebesar 0.73 pada bulan januari dan 0,93 pada bulan februari tahun 2020.

Kata kunci— Demam Berdarah Dengue, DBD, *K-means*, *Correlation matrix*, Machine learning.

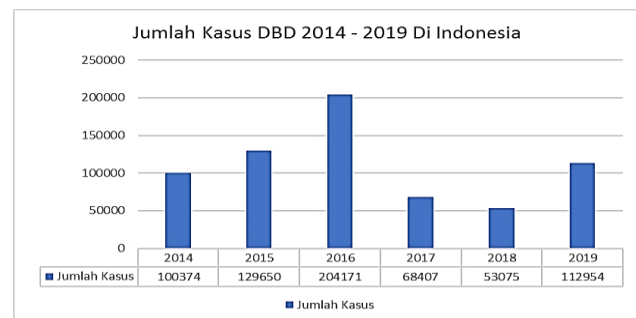
Abstract— *Dengue cases in Pekanbaru in November 2020 reached 2,788 cases and 33 deaths. The government has carried out socialization to eradicate mosquito nests and provided vector control tools and materials. However, the government's efforts were not practical because the applied method has not been able to refer to vector data and information. Machine learning can be used to analyze specific problems such as Dengue. Therefore this study employed a Machine Learning algorithm, i.e., k-means clustering and correlation matrix for dengue risk analysis in Pekanbaru. This study obtained 12 sub-districts and 50 dengue attributes and weather data in 2020. K-means automatically searches for unknown clusters from dengue cases data quickly, which cluster results C1 (Sukajadi, Senapelan), C2 (Tenayan Raya, Tampan), and C3 (Rumbai Pesisir, Rumbai). Based on experimental testing, this study produced a silhouette score is 0.6. Meanwhile, the correlation matrix looks for relevant relationships hidden in the data. The correlation matrix*

obtained a strong linear relationship between the population (JP) and sufferers (P) of 0.73 for January and 0.93 for February 2020.

Keywords— *Dengue Fever, K-means, Correlation matrix, Machine learning.*

I. PENDAHULUAN

Demam Berdarah *Dengue* (DBD) memerlukan perhatian dan penanganan yang tepat di Indonesia. Berdasarkan data kasus demam berdarah di Indonesia pada tahun 2014 – 2019, demam berdarah masih menjadi penyakit menular yang terjadi setiap tahunnya. Data jumlah kasus DBD dapat dilihat pada Gambar 1 berikut :



Gambar. 1 Grafik jumlah kasus DBD 2014 – 2019 di Indonesia

Provinsi Riau Merupakan salah satu daerah endemis DBD di Indonesia dengan 2.788 kasus dan 33 jumlah kematian pada tahun 2020 [1]. Kota Pekanbaru merupakan kasus demam berdarah terbanyak kedua di Provinsi Riau dengan 474 kasus dan 4 kematian [2]. Hal ini membuat pemerintah harus mencari mitigasi yang tepat untuk menekan angka kasus DBD di Kota Pekanbaru.

Pemerintah telah berupaya dalam menekan angka DBD dengan melakukan sosialisasi pemberantasan sarang nyamuk (PSN) dan memberikan dukungan bahan dan alat pengendalian vektor [3]. Namun, upaya pemerintah tersebut masih belum efektif, terlihat dari data DBD yang tidak jauh berubah setiap tahunnya.

Upaya pengendalian vektor yang dilakukan hampir di semua negara yang merupakan negara endemis masih tidak tepat sasaran, tidak berkesinambungan dan belum memutuskan rantai penularan [4]. Hal ini disebabkan oleh metode yang diterapkan dianggap belum mampu mengacu pada data dan informasi vektor. Oleh karena itu, *Machine Learning* yang menggabungkan informatika dan analisis statistik dapat digunakan untuk menganalisis masalah yang khusus seperti DBD.

Penelitian terkait penggunaan *machine learning* untuk kasus DBD menggunakan *k-means clustering* pernah dilakukan oleh Hariyanto pada tahun 2019 [5]. Pada penelitian tersebut membagi daerah menjadi 3 *cluster* yaitu endemis (C1), potensi (C2), dan sporadis (C3). Namun, pada penelitian tersebut tidak adanya analisis risiko terhadap hubungan antar atribut di dalam data. Sedangkan, pengetahuan tentang hubungan antar atribut penting untuk diketahui dalam menentukan kebijakan yang tepat dalam pencegahan dan mitigasi penyebaran DBD.

Berdasarkan tinjauan pustaka, algoritma *k-means* mampu mengelompokkan daerah penyebaran DBD berdasarkan karakteristik data yang mirip. Analisis hubungan atribut menggunakan *Correlation matrix* pernah dilakukan oleh Kurniawan pada tahun 2020, dalam menganalisis risiko penyebaran Covid-19 pada 200 negara. Pada penelitian tersebut menunjukkan adanya korelasi linear positif yang kuat yaitu 0,78 dari jumlah kasus dengan total kematian dan 0,85 dari total kematian dengan dikategorikan kritis [6].

Oleh karena itu, pada penelitian ini menghibridisasi algoritma *k-means clustering* dan *correlation matrix* untuk menganalisis risiko penyebaran DBD di Kota Pekanbaru. Penelitian ini diharapkan mampu mengelompokkan Kecamatan di Kota Pekanbaru dengan karakteristik yang mirip. Sehingga dapat menjadi rujukan bagi pemerintah dalam upaya pencegahan dan mitigasi yang tepat untuk setiap daerah rawan DBD. Selain itu, penelitian ini diharapkan mampu menentukan atribut yang berpengaruh terhadap angka DBD di Pekanbaru.

A. Demam Berdarah Dengue (DBD)

Demam Berdarah *Dengue* merupakan penyakit yang disebabkan oleh virus *Dengue*. Penyakit Demam berdarah ini ditularkan melalui gigitan nyamuk *Aedes aegypti*, *Aedes albopictus*, dan *Aedes polynesiensis* kepada manusia [7].

Pemerintah diharapkan melakukan pemberantasan DBD lebih di fokuskan kepada kecamatan yang memiliki angka kepadatan penduduk tinggi. Selain itu, adanya hubungan kausal antara curah hujan, kelembaban udara, dan suhu terdapat kasus DBD [8]. Oleh sebab itu pengetahuan tentang data serta analisa yang tepat mengenai cuaca di suatu daerah sangat diperlukan untuk meningkatkan kewaspadaan dini terhadap kasus DBD.

B. Metode K-Means Clustering

Pada dasarnya *K-Means* menggunakan beberapa komponen yang diperoleh sebagai pusat *cluster* awal, pusat *cluster* ditentukan dengan memilih dari *dataset* secara acak. Setelah itu *k-means* akan memeriksa tiap komponen dari *dataset*, lalu memberi tanda tiap komponen ke salah satu pusat *cluster* yang diberi label atau didefinisikan sebelumnya. Selanjutnya akan dihitung kembali untuk menentukan pusat *cluster* kemudian akan terbentuk pusat *cluster* baru. Proses ini selesai hingga pusat *cluster* tidak berubah dengan pusat *cluster* sebelumnya [9].

Untuk menentukan pusat *cluster* terdekat dari setiap *dataset* menggunakan persamaan *Eucliden distance* :

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (1)$$

Keterangan :

D_e : *Euclidean Distance*

i : Banyaknya objek,

(x,y) : merupakan koordinat objek

(s,t) : merupakan koordinat *centorid* (titik pusat *cluster*).

Memperbaharui titik pusat *cluster* dengan persamaan titik pusat *cluster* :

$$\bar{v}_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \quad (2)$$

Keterangan :

\bar{v}_{ij} : centroid/rata-rata *cluster* ke- i untuk *variable* ke- j

N_i : Jumlah data yang menjadi anggota *cluster* ke- i

i,k : indeks dari *cluster*

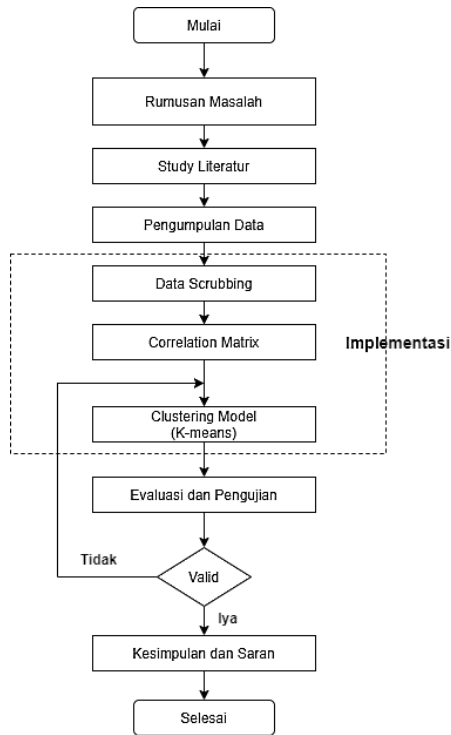
C. Correlation Matrix

Matriks korelasi adalah matrik yang menunjukkan koefisien korelasi antara kumpulan atribut. Sebuah survei menyatakan bahwa pengetahuan yang dihasilkan *correlation matrix* sangat penting. Setiap atribut acak (X_i) dalam tabel berkorelasi dengan masing-masing nilai lain dalam tabel (X_j). Nilai besar dalam matriks ini menunjukkan hubungan yang serius antara atribut yang terlibat. Ini memungkinkan Anda untuk melihat pasangan mana yang memiliki korelasi tertinggi.

Matriks korelasi menjadikan semua atribut dengan dasar yang sama dalam hal ini setiap analisis dapat digambarkan sebagai analisis saling ketergantungan. Namun terkadang, atribut tidak memiliki status yang sama. Dalam kasus seperti itu, mungkin peneliti dapat melihat pada bagaimana beberapa atribut bergantung pada yang lain dan mungkin tertarik untuk mengetahui bagaimana urutan berikutnya bergantung pada yang datang lebih awal.

II. METODOLOGI PENELITIAN

Metodologi penelitian dilakukan untuk mendapatkan hasil sesuai dengan yang diinginkan. Berikut tahap-tahapan dalam penelitian ini dapat dilihat pada Gambar 2 berikut :



Gambar. 2 Desain penelitian

A. Metode Pengumpulan Data

Metode yang dilakukan dalam mengumpulkan data dan informasi mengenai penelitian yang dilakukan diperoleh dari *study literature*, Dinas Kesehatan Pekanbaru dan Badan Meteorologi, Klimatologi, dan Geofisika (BMKG Kota Pekanbaru). Data yang digunakan yaitu data DBD per Kecamatan Kota Pekanbaru tahun 2020 dan data cuaca Pekanbaru 2020. Data dapat dilihat pada Tabel I dan II berikut :

TABEL I
DATA DBD PER KECAMATAN KOTA PEKANBARU TAHUN 2020

Kecamatan	P1	M1	IR1	CFR1	P2	...	CFR12	JP	KP
Sukajadi	7	0	1.63	0	4	...	0	42852	11396.81
Senapelan	8	0	2.26	0	4	...	0	35357	5316.84
Pekanbaru Kota	1	0	0.44	0	2	...	0	22604	10001.77
Rumbai Pesisir	4	0	0.57	0	5	...	0	70488	448.03
Rumbai	10	0	1.28	0	5	...	0	78185	606.79
Lima Puluh	14	0	3.63	0	6	...	0	38613	9557.67
Sail	0	0	0	0	1	...	0	20384	6252.76
Bukit Raya	8	0	0.86	0	12	...	0	93478	4239.37
Marpoyan Damai	10	0	0.78	0	15	...	0	127600	4290.52
Tenayan Raya	27	0	1.75	0	15	...	0	154261	900.69
Tampan	16	0	0.79	0	17	...	0	203238	3398.06
Payung Sekaki	11	0	1.14	0	12	...	0	96296	2227.01

TABEL II
DATA CUACA PEKANBARU TAHUN 2020

Bln	P	M	IR	CFR	Tn	Tx	Tavg	RH_avg	RR	ss
Jan	116	0	27.07	0	23.8	31.94	27.2	86	261	117.2
Feb	98	1	27.72	0.28	23.73	32.52	27.3	79	199	126.7
Mar	82	0	36.28	0	24.08	33.4	27.8	79	182	135.7
Apr	63	0	8.94	0	24.04	33.24	27.8	83	335	118.4
Mei	26	0	3.33	0	24.25	33.45	28	80	223	149.3
Jun	32	0	8.29	0	23.68	32.4	27.2	81	107	141.9
Jul	26	1	12.76	0.49	23.35	32.5	26.8	81	229	149.3
Ags	15	1	1.6	0.11	23.5	33.28	27.5	77	47	137.9
Spt	9	0	0.71	0	23.04	32.79	26.7	81	292	80.8
Okt	7	0	0.45	0	23.34	33.21	27.3	79	161	132.8
Nov	12	0	0.59	0	23.45	32.35	26.6	85	512	100
Des	15	0	1.56	0	23.25	32.08	26.7	81	190	133.7

Keterangan:

- P : Penderita
- P1 : Penderita pada bulan 1
- M : Kasus meninggal
- M1 : Kasus meninggal pada bulan 1
- IR : *Incident Rate*
- CFR : *Case Fatality Rate*
- Tn : Temperatur minimum (°C)
- Tx : Temperatur maksimum (°C)
- Tavg : Temperatur rata-rata (°C)
- RH_avg : Kelembapan rata-rata (%)
- RR : Curah hujan (mm)
- ss : Lamanya penyinaran matahari (jam)
- JP : Jumlah penduduk
- KP : Kepadatan penduduk

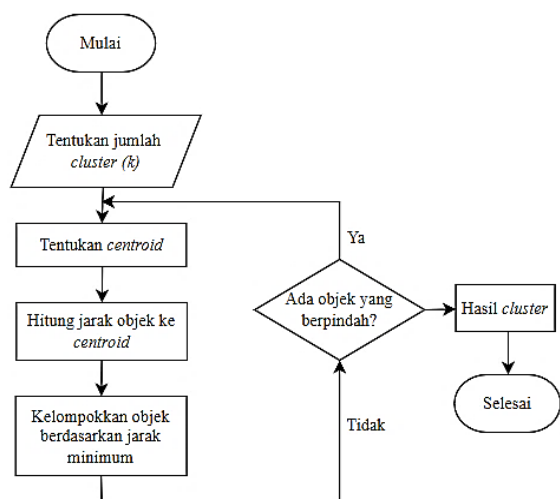
B. Implementasi

Implementasi pada penelitian ini menggunakan bahasa pemrograman *python*, dengan tiga tahap yaitu :

1) *Data Scrubbing* : Merupakan proses pembersihan *missing value* untuk digunakan pada saat pemodelan, sehingga data lebih mudah dikenali dan dapat meningkatkan proses pembelajaran dari algoritma *machine learning* yang digunakan.

2) *Correlation Matrix* : Pada tahap ini, data yang sudah di bersihkan pada tahap sebelumnya kemudian akan di lakukan proses matriks korelasi untuk menunjukkan keterkaitan antara kumpulan atribut. Setiap atribut acak (Xi) dalam tabel berkorelasi dengan masing-masing nilai lain dalam tabel (Xj). Nilai besar dalam matriks ini menunjukkan kolinearitas yang serius antara atribut yang terlibat. Ini memungkinkan untuk melihat pasangan antar atribut yang memiliki korelasi tertinggi.

3) *Clustering Model* : Merupakan tahap implementasi algoritma *k-means* atau pemodelan data dengan *k-means*. Tahap implementasi *clustering* daerah rawan DBD menggunakan *k-means* dapat dilihat dalam *flowchart* pada Gambar 3 berikut :



Gambar. 3 Flowchart *k-means*

Berikut Tahap *clustering* menggunakan *k-means* berdasarkan Gambar 3 :

1) *Menentukan jumlah cluster (k)* : Tahap awal menentukan jumlah *cluster*, pada penelitian ini menggunakan 3 *cluster*.

2) *Menentukan Centroid* : Mengambil satu pusat data yang mewakili suatu *cluster*.

3) *Menghitung jarak objek ke centroid* : Menghitung jarak objek ke *centroid* dengan menggunakan persamaan *euclidian distance* pada formula (1).

4) *Mengelompokkan objek berdasarkan jarak minimum* : Langkah selanjutnya mengelompokkan objek ke dalam *cluster* berdasarkan jarak minimum terhadap *centroid*. Sehingga didapatkan 3 *cluster* dengan beberapa objek sebagai anggota *cluster*.

5) *Memperhatikan perpindahan objek* : Jika objek mengalami perpindahan *cluster* maka proses diulang dari tahap 2. Namun jika tidak terdapat perpindahan objek terhadap *cluster*, maka iterasi selesai.

Tahap *clustering* menggunakan algoritma *k-means* ditampilkan dalam *pseudocode* dibawah :

```

Program Algoritma K-Means
Input:
  D= {t1,t2,...Tn} //Set of elements
  K      //Number of desired custer
Output:
  K      //set of clusters
K-Means algorithm :
  Assign initial values for m1,m2,...mk
  Repeat
  Assign each item ti to the clusters which
  has the closest mean;
  Calculate new mean for each cluster;
  Until
  Convergence criteria is met;
  
```

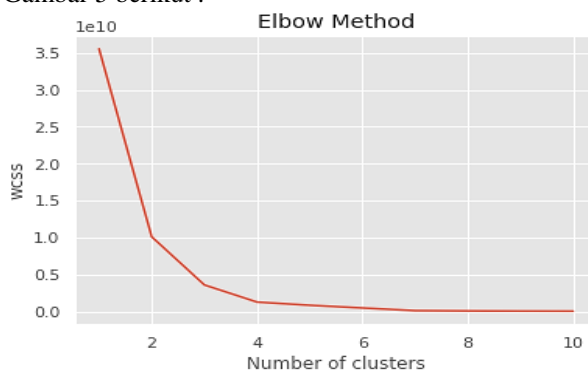
C. Evaluasi dan Pengujian

Pada tahap ini, akan dilakukan evaluasi dan pengujian terhadap kualitas dan efektivitas model yang dibuat dengan

algoritma *k-means* dan *correlation matrix*. Pengujian dilakukan menggunakan *Silhouette Coefficient*. Tujuan dari pengujian ini yaitu untuk melihat kualitas dan kekuatan *cluster* yang diperoleh dari suatu *clustering* data.

III. HASIL DAN PEMBAHASAN

Pada penelitian ini dilakukan tiga eksperimen menggunakan data yang telah diperoleh dengan menggunakan algoritma *K-Means* dan *Correlation Matrix*. Eksperimen pertama yaitu dengan mengelompokkan berdasarkan kecamatan pada tahun 2020, eksperimen kedua yaitu mengelompokkan berdasarkan kecamatan setiap bulan pada tahun 2020 terakhir eksperimen ketiga menentukan hubungan terkuat antar atribut yang berpengaruh terhadap peningkatan DBD di Kota Pekanbaru dengan menggunakan *corellation matrix*. Penentuan daerah rawan DBD menjadi tiga kategori yang merupakan jumlah *cluster (k)* terbaik, ditampilkan menggunakan metode *elbow* yang dapat dilihat pada Gambar 5 berikut :



Gambar. 4 Penentuan jumlah *cluster* menggunakan metode *elbow*

Berdasarkan metode *elbow*, *number of cluster* terlihat penurunan membentuk siku pada $k=3$, hal ini menunjukkan bahwa pembagian daerah rawan DBD berdasarkan 3 kategori merupakan nilai *cluster* terbaik.

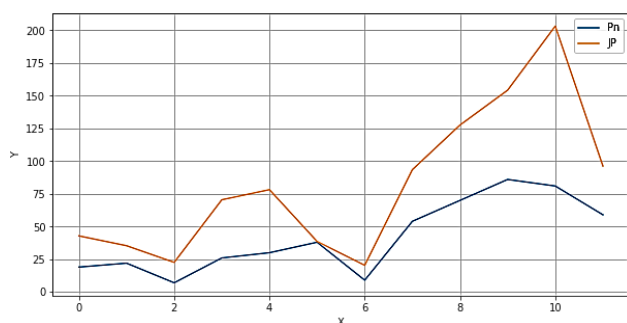
A. Pengelompokkan Berdasarkan Kecamatan Pada Tahun 2020

Berdasarkan hasil *clustering* 12 kecamatan di Kota Pekanbaru menggunakan data pada Tabel I didapatkan hasil:

- Daerah *cluster 1* : Sukajadi, Senapelan, Pekanbaru Kota, Lima Puluh, dan Sail
- Daerah *cluster 2* : Tenayan Raya dan Tampan
- Daerah *cluster 3* : Rumbai Pesisir, Rumbai, Bukit Raya, Marpoyan Damai, dan Payung Sekaki.

Visualisasi hasil *Cluster* Kecamatan dapat dilihat pada Gambar 6 berikut :

Berdasarkan tabel IV di atas menunjukkan adanya korelasi linear positif yang kuat antara Penderita (P) dengan jumlah penduduk (JP), yaitu sebesar 0.73 pada bulan 1 dan 0,93 pada bulan 2, begitu pula dengan korelasi pada bulan selanjutnya. Nilai positif membuktikan arah korelasi, yaitu jika salah satu atribut meningkat, maka atribut yang lain juga meningkat. Hubungan antara penderita dan jumlah penduduk dalam satu tahun dapat dilihat pada Gambar 7 berikut:



Gambar. 7 Grafik hubungan penderita (P) dengan jumlah penduduk (JP)

Keterangan :

- Pn : Penderita total dalam satu tahun
- JP : Jumlah penduduk

Berdasarkan Gambar 9 menunjukkan bahwa fitur Penderita (P) dengan Jumlah Penduduk (JP) mempunyai bentuk garis yang mirip. Hal ini menunjukkan bahwa, semakin tinggi jumlah penduduk maka akan semakin tinggi berisiko terkena DBD. Selain itu, dapat menjadi rujukan bagi pihak terkait dalam penanganan dan mitigasi yang tepat untuk menekan angka kasus DBD di Pekanbaru.

IV. KESIMPULAN

Berdasarkan hasil penelitian dalam menganalisis risiko penyebaran DBD di Kota Pekanbaru menggunakan algoritma *k-means clustering* dan *correlation matrix*, dapat disimpulkan bahwa penelitian ini berhasil mengelompokkan daerah DBD di Kota Pekanbaru secara cepat. Berdasarkan pengujian didapatkan silhouette score 0,6. Kecamatan Tenayan Raya merupakan daerah dengan angka kasus tertinggi yang masuk pada *cluster 2*. Kecamatan yang berada pada *cluster 1* sangat berisiko menjadi *cluster 2* pada bulan berikutnya. Berdasarkan hasil *corellation matrix* menggunakan data cuaca, didapatkan korelasi linear positif pada atribut penderita (P) dan suhu rata-rata minimum (Tx) dengan angka 0,57. Sedangkan *corellation matrix* menggunakan data DBD, didapatkan korelasi linear positif pada atribut penderita (P) dan jumlah

penduduk (JP) sebesar 0,73 untuk bulan pertama dan 0,93 untuk bulan kedua, dan begitu pula dengan bulan selanjutnya yang menghasilkan korelasi tinggi. Hal ini membuktikan adanya korelasi yang kuat antara jumlah penduduk dengan jumlah penderita.

UCAPAN TERIMA KASIH

Terima kasih kepada Dinas Kesehatan Kota Pekanbaru, BMKG Stasiun Klimatologi Kampar-Riau, dan lembaga riset CIDSCI (*Center of Islamic Data Science and Continuous Improvement*) UIN Suska Riau yang telah mendukung penelitian ini.

REFERENSI

- [1] H. Riau, "Hingga Oktober 2020, Sudah 2.788 Warga Riau Kena DBD," *www.halloriau.com*, 2020. <https://www.halloriau.com/read-otonomi-139242-2020-11-26-hingga-oktober-2020-sudah-2788-warga-riau-kena-dbd.html> (accessed Apr. 22, 2021).
- [2] B. Pemko, "Angka DBD di Kota Pekanbaru Mencapai 474 Kasus," *www.pekanbaru.go.id*, 2020. <https://www.pekanbaru.go.id/p/news/angka-dbd-di-kota-pekanbaru-mencapai-474-kasus> (accessed Apr. 23, 2021).
- [3] K. K. R. Indonesia, "Kesiapsiagaan Menghadapi Peningkatan Kejadian Demam Berdarah Dengue Tahun 2019," *p2p.kemkes.go.id*, 2019. <http://p2p.kemkes.go.id/kesiapsiagaan-menghadapi-peningkatan-kejadian-demam-berdarah-dengue-tahun-2019/> (accessed Apr. 25, 2021).
- [4] N. Susianti, "Strategi Pemerintah Dalam Pemberantasan Demam Berdarah Dengue (Dbd) Di Kabupaten Merangin," *Bul. Penelit. Sist. Kesehat.*, vol. 22, no. 1, pp. 34–43, 2019, doi: 10.22435/hsr.v22i1.1799.
- [5] M. Hariyanto and R. T. Shita, "Clustering pada Data Mining untuk Mengetahui Potensi Penyebaran Penyakit DBD Menggunakan Metode Algoritma K-Means dan Metode Perhitungan Jarak Euclidean Distance," *Sist. Komput. dan Tek. Inform.*, vol. 1, no. 1, pp. 117–122, 2018.
- [6] R. Kurniawan, S. N. H. S. Abdullah, F. Lestari, M. Z. A. Nazri, A. Mujahidin, and N. Adnan, "Clustering and Correlation Methods for Predicting Coronavirus COVID-19 Risk Analysis in Pandemic Countries," *2020 8th Int. Conf. Cyber IT Serv. Manag. CITSM 2020*, 2020, doi: 10.1109/CITSM50537.2020.9268920.
- [7] K. Fatmawati and A. P. Windarto, "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi," *Comput. Eng. Sci. Syst. J.*, vol. 3, no. 2, p. 173, 2018, doi: 10.24114/cess.v3i2.9661.
- [8] A. S. Ichwani and H. A. Wibawa, "Prediksi Angka Kejadian Demam Berdarah Dengue (DBD) Berdasarkan Faktor Cuaca Menggunakan Metode Extreme Learning Machine (Studi Kasus Kecamatan Tembalang)," *J. IPTEK*, vol. 23, no. 1, pp. 31–38, 2019, doi: 10.31284/j.iptek.2019.v23i1.471.
- [9] R. K. Dinata, S. Safwandi, N. Hasdyna, and N. Azizah, "Analisis K-Means Clustering pada Data Sepeda Motor," *INFORMAL Informatics J.*, vol. 5, no. 1, p. 10, 2020, doi: 10.19184/isj.v5i1.17071.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

