Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK) DOI: 10.25126/jtiik.1148703
Vol. 1, No. 4, Agustus 2024, hlm. 877-886 p-ISSN: 2355-7699
Akreditasi KEMENRISTEKDIKTI, No. 36/E/KPT/2019 e-ISSN: 2528-6579

PENERAPAN METODE *K-MEANS CLUSTERING* DAN *SIMPLE MOVING AVERAGE* UNTUK MEMPREDIKSI JENIS PENYAKIT DI PROVINSI JAWA TIMUR

Shynta Ayu Dwi Darmawan[1], Karmilasari[*2]

[1,2]Universitas Gunadarma, Depok
Email: [1]shynta.add@gmail.com, [2]karmila@staff.gunadarma.ac.id
[*]Penulis Korespondensi

(Naskah masuk: 05 Februari, diterima untuk diterbitkan: 12 Agustus 2024)

Abstrak

Berdasarkan buku Profil Kesehatan Provinsi Jawa Timur dari tahun 2017 hingga 2021, terlihat bahwa banyak penyakit menjadi perhatian di 38 kota atau kabupaten di Jawa Timur. Dari data tersebut, belum jelas daerah mana yang perlu mendapatkan prioritas penanganan. Terdapat kesulitan pengidentifikasian daerah prioritas dalam penanganan penyakit. Tujuan penelitian adalah mengelompokkan kasus ke dalam kelompok yang relevan dan dapat diidentifikasi, memprediksi tren kasus penyakit berdasarkan data historis di setiap daerah dari tahun ke tahun, membangun sistem berbasis *website* sebagai media implementasi prediksi dan *clustering*. Tahapan penelitian meliputi studi literatur, pengembangan model *clustering* menggunakan *K-Means* dan prediksi menggunakan *Simple Moving Average*, pengembangan sistem menggunakan *MySQL*, *PHP* dan *Angular*. Metode *K-Means Clustering* membagi data ke dalam beberapa kelompok berdasarkan karakteristik yang mirip, sehingga lebih mudah untuk mengidentifikasi pola dan tren yang tersembunyi dalam data kesehatan. *Simple Moving Average* menggunakan rata-rata data masa lalu untuk memperhalus perubahan jangka pendek dan menemukan pola jangka panjang. Hasil penelitian menunjukkan berhasilnya implementasi *clustering* dalam 3 *cluster*: rendah, sedang dan tinggi. Prediksi tren kasus penyakit berhasil diterapkan menggunakan data historis periode 5 tahun, memberikan wawasan signifikan untuk perencanaan kesehatan. Uji fungsional dan kompatibilitas *browser* menunjukkan bahwa sistem berjalan sesuai harapan di berbagai lingkungan. Pengujian *usability* dengan metode *WebQual* 4.0 menunjukkan nilai rata-rata 4,34 (sangat baik), mengonfirmasi keberhasilan sistem dalam memenuhi kebutuhan pengguna. Hasil penelitian ini memiliki implikasi praktis untuk meningkatkan efektivitas pengelolaan penanggulangan penyakit di Provinsi Jawa Timur dengan mengidentifikasi prioritas daerah secara lebih akurat dan membuat keputusan yang lebih baik berdasarkan data.

Kata kunci: *clustering, k-means, prediksi, Simple Moving Average, jumlah kasus penyakit, dinas kesehatan provinsi jawa timur*

IMPLEMENTATION OF K-MEANS CLUSTERING AND SIMPLE MOVING AVERAGE METHODS TO PREDICT DISEASE TYPES IN EAST JAVA PROVINCE

Abstract

Based on the East Java Provincial Health Profile book from 2017 to 2021, it appears that many diseases are a concern in 38 cities or districts in East Java. From this data, it is not clear which areas need priority handling. There are difficulties in identifying priority areas in disease management. The research objectives are to cluster cases into relevant and identifiable groups, predict trends in disease cases based on historical data in each region from year to year, build a website-based system as a medium for implementing predictions and clustering. The research stages include literature study, clustering model development using K-Means and prediction using Simple Moving Average, system development using MySQL, PHP and Angular. The K-Means Clustering method divides data into groups based on similar characteristics, making it easier to identify patterns and trends hidden in health data. Simple Moving Average uses the average of past data to smooth out short-term changes and find long-term patterns. The results showed the successful implementation of clustering in 3 clusters: low, medium and high. Disease case trend prediction was successfully implemented using historical data over a 5-year period, providing significant insights for health planning. Functional and browser compatibility tests showed that the system runs as expected in various environments. Usability testing with the WebQual 4.0 method showed an average score of 4.34 (excellent), confirming the success of the system in meeting user needs. The results of this study have practical

implications for improving the effectiveness of disease management in East Java Province by more accurately identifying regional priorities and making better decisions based on data.

Keywords: *clustering, k-means, prediction, Simple Moving Average, number of disease cases, east java provincial health office*

1. PENDAHULUAN

Penanganan kesehatan merupakan tanggung jawab semua pihak termasuk pemerintah di tingkat provinsi. Pemerintah provinsi berperan dalam memantau kualitas layanan dan fasilitas kesehatan yang disediakan di tingkat provinsi dan memastikan bahwa layanan tersebut mematuhi pedoman dan standar nasional. Dinas Kesehatan Provinsi Jawa Timur bertanggung jawab membawahi 38 Dinas Kesehatan Kota atau Kabupaten di wilayah tersebut. Salah satu tanggung jawab Dinas Kesehatan memantau kasus penyakit yang ada di wilayahnya. Ada sekitar 13 jenis penyakit sesuai dengan ketentuan dari Kementrian Kesehatan RI yang harus selalu dipantau dan dilaporkan. Saat ini publikasi jumlah kasus setiap penyakit setiap tahun di provinsi Jawa Timur masih terbatas pada laporan Profil Kesehatan yang dikeluarkan oleh Dinas Kesehatan Jawa Timur dalam bentuk *PDF* (*Portable Document Format*). Hanya dapat diakses oleh kalangan terbatas, sehingga *stakeholder* lain atau masyarakat luas tidak mengetahui adanya publikasi tersebut. Padahal di sisi lain informasi mengenai kasus penyakit di suatu daerah penting untuk diketahui secara luas sehingga penanganannya menjadi lebih komprehensif dan melibatkan berbagai pihak.

Pada masa keterbukaan informasi saat ini, informasi mengenai jumlah kasus penyakit di suatu daerah seharusnya dapat diakses secara luas dengan memanfaatkan teknologi informasi. Penelitian ini bertujuan untuk mengembangkan sistem informasi sebagai media implementasi pengelompokan dan prediksi jumlah kasus penyakit tiap kota atau kabupaten di Jawa Timur.

Beberapa penelitian sejenis sebelumnya di antaranya dilakukan oleh Tanty, dkk yang melakukan penelitian mengenai pengelompokan penyakit pada pasien berdasarkan usia. Tapi kurang adanya analisis tren waktu (Tanty, Serasi Ginting dan Simanjuntak, 2021). Sementara itu (Sembiring, 2022) berhasil mengelompokan wilayah penyebaran demam berdarah *dengue* (DBD) tapi tidak memberikan prediksi jangka panjang. Penelitian (Adiputra, 2022) di RS Dharma Kerti terbatas pada tingkat rumah sakit dan tidak memiliki cakupan yang lebih luas. Penelitian (Sulistiyo Hidayat et al., 2022) mengelompokkan kasus Tuberkulosis tanpa memprediksi tren jangka panjang.

Sedangkan (Purwanto, Nilogiri dan Wardoyo, 2022) mengelompokkan penyebaran *TBC* di puskesmas Kabupaten Jember, tetapi tidak dapat diterapkan di wilayah yang lebih luas. Penelitian (Firdaus, Aisyiyah dan Devi, 2022) mengelompokan penyakit di posyandu lansia, meskipun terbatas pada satu kelompok usia. Mengelompokkan lokasi penanganan diare pada balita, menunjukkan manfaat dalam menargetkan perawatan kesehatan tetapi kurang dalam analisis data jangka panjang (Syamfithriani, Mirantika dan Trisudarmo, 2023). Pengelompokan *tweet* tentang *Covid-19* untuk memberikan perspektif media sosial, namun terbatas pada interpretasi teks (Dani, Julidany dan David, 2023).

Beberapa penelitian telah berhasil menggunakan metode *K-Means*, hanya terbatas pada pengelompokan dan tidak memberikan analisis atau prediksi keberlanjutan tambahan. Pengelompokan pengeluaran rumah sakit dengan BPJS (Nabyla dan Mukrodin, 2021). Pengelompokan pasien BPJS di Rumah Sakit Anwar Medika Balong Bendo Sidoarjo (Ali dan Masyfufah, 2021). Pengelompokan kasus balita berpotensi mengalami gizi buruk (Rani dan Novianti, 2022). Pengelompokan persediaan obat di apotek (Ayu ramadhanty et al., 2022).

Selanjutnya pengelompokan obat di RS (Nugroho, Hendrawan dan Puwantoro, 2022). Pengelompokan penyebaran penyakit *ISPA* di Riau (Purba, Poningsih dan Tambunan, 2021), dan di Pematangsiantar (Agoestina, Tambunan dan Nasution, 2022). Pengelompokan data pasien penyakit jantung (Al Rivan dan Sonaru, 2022). Pengelompokan status *Covid-19* oleh Tambunan, dkk (Tambunan et al., 2022).

Berikutnya pengelompokan status daerah terdampak *Covid-19* (Juliana, Aleyda dan Yuliana, 2021). Pengelompokan penyakit menular (Tampubolon dan Handoko, 2021). Pengelompokan rata rata konsumsi kalori dan protein berdasarkan provinsi (Aryadi, Danreas dan Sari, 2022). Pengelompokan penyakit di RS (Nurhidayah, Lestari dan Purwanto, 2022). Pengelompokan daerah rawan kasus penyakit di Kabupaten Sleman (Rokhimah dan Kismiantini, 2022). Pengelompokan penularan penyakit *HIV/AIDS* (Sari et al., 2023).

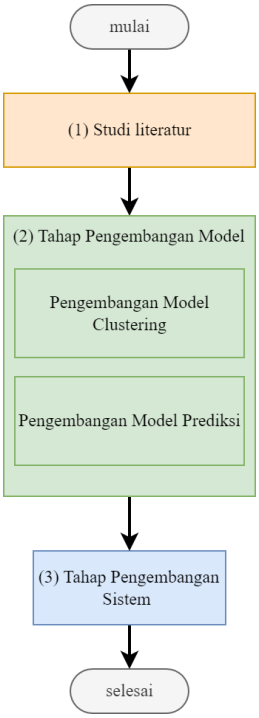
Penelitian ini memberikan berbagai kontribusi yang berbeda. Pertama, dengan mengelompokkan penyakit di tingkat provinsi menggunakan data historis yang lebih panjang, analisis yang lebih mendalam dan bermakna dapat dilakukan. Kedua, kombinasi metode *K-Means Clustering* dan *Simple Moving Average* memberikan strategi komprehensif yang tidak hanya mengelompokkan data penyakit, tetapi juga meramalkan tren kasus penyakit dengan menggunakan data historis.

Diharapkan melalui penelitian ini Dinas Kesehatan Provinsi maupun Kota Kabupaten dan

stakeholder terkait serta masyarakat luas dapat memantau jumlah kasus penyakit di daerah masing-masing. Selanjutnya memberikan informasi yang akurat dan komprehensif mengenai prediksi tren penyakit di Jawa Timur, sehingga dapat membantu pengambilan keputusan yang lebih baik dalam pengelolaan kesehatan.

2. METODE PENELITIAN

Pelaksanaan penelitian dibagi menjadi 3 tahapan yang disajikan dalam bentuk bagan air pada Gambar 1.



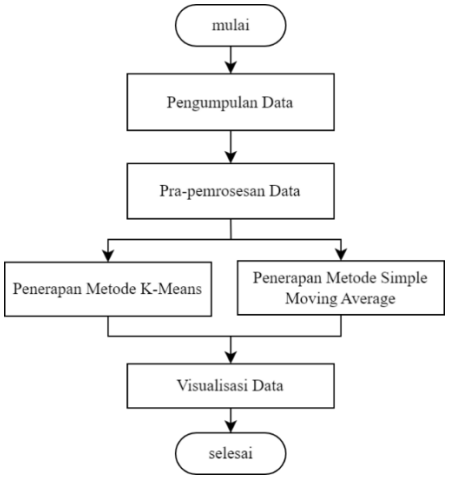
Gambar 1 Tahapan Penelitian

2.1 Studi Literatur

Penerapan metode *K-Means Cluster*ing mengelompokkan data berdasarkan karakteristik yang mirip, sehingga lebih mudah untuk menemukan pola dan tren yang tersembunyi dalam data kesehatan (Utari, 2021). Efisiensi metode dalam menangani set data yang sangat besar, serta menghasilkan output yang mudah diinterpretasikan, mendorong pengambilan keputusan yang lebih tepat dalam bidang kesehatan (Andi, Juliandy and David, 2023). Sedangkan *Simple Moving Average* adalah metode prediksi menggunakan rata-rata data sebelumnya untuk memperhalus perubahan jangka pendek sambil mengidentifikasi pola jangka panjang (Aji et al., 2022). Kelebihannya adalah jumlah titik data dalam setiap rata-rata tetap konstan dari waktu ke waktu.

2.2 Tahap Pengembangan Model

Dalam penelitian, pembuatan model dibagi menjadi 2 tahap yang saling berhubungan: pengembangan model pengelompokan menggunakan metode *K-Means* dan pengembangan model prediksi menggunakan metode *Simple Moving Average* (*SMA*). Berikut adalah bagan air yang menggambarkan alur pengembangan model disajikan pada Gambar 2.



Gambar 2 Tahap Pengembangan Model

Pengumpulan Data

Proses pengumpulan data dari Buku Profil Kesehatan Provinsi Jawa Timur berbentuk *PDF* (*Portable Document Format*), yang diterbitkan oleh Dinas Kesehatan Provinsi Jawa Timur dan dapat diakses melalui situs web resmi di https://dinkes.jatimprov.go.id, dari tahun 2017 hingga tahun 2021.

Pra-pemrosesan Data

Data harus diambil dari *file PDF*. Berikut tampilan data jumlah kasus penyakit tuberkulosis dalam bentuk *PDF* disajikan pada Gambar 3.

Selanjutnya, dilakukan dengan men-*screenshot* halaman-halaman yang berisi tabel atau informasi yang ingin diambil. Kemudian mengubah *file* gambar tersebut ke dalam format tabel *Excel*, dapat memanfaatkan alat konversi *online* https://nanonets.com/image-to-excel. Kemudian mengambil data yang diperlukan sehingga diperoleh hasil akhir berupa data bersih, dapat dilihat pada Tabel 1.

Tabel 1 Data Bersih Sebagai Bahan *Clustering* dan Prediksi

Nama Daerah	2017	2018	2019	2020	2021
KAB. PACITAN	214	611	362	288	212
KAB. PONOROGO	977	1065	1158	904	687
KAB. TRENGGALEK	401	465	552	402	239
KAB. TULUNGAGUNG	1043	1188	1236	821	716
KAB. BLITAR	462	695	853	641	485
…	…	…	…	…	…
KOTA PASURUAN	499	569	581	521	584
KOTA MOJOKERTO	497	391	455	330	482
KOTA MADIUN	634	692	709	526	447
KOTA SURABAYA	6338	7007	7950	4151	4631
KOTA BATU	32	200	241	185	140

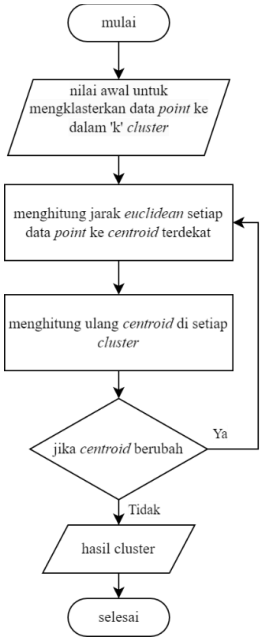
JUMLAH TERDUGA TUBERKULOSIS, KASUS TUBERKULOSIS, KASUS TUBERKULOSIS ANAK, CASE NOTIFICATION RATE (CNR) PER 100.000 PENDUDUK DAN CASE DETECTION RATE (CDR) MENURUT JENIS KELAMIN
PROVINSI JAWA TIMUR
TAHUN 2021

NO	KABUPATEN/KOTA	PUSKESMAS	JUMLAH TERDUGA TUBERKULOSIS YANG MENDAPATKAN PELAYANAN SESUAI STANDAR	JUMLAH SEMUA KASUS TUBERKULOSIS					KASUS TUBERKULOS ANAK 0-14 TAH
				LAKI-LAKI		PEREMPUAN		LAKI-LAKI + PEREMPUAN	
				JUMLAH	%	JUMLAH	%		
1	2	3	4	5	6	7	8	9	10
1	KAB. PACITAN	24	2.290	118	55,7	94	44,3	212	
2	KAB. PONOROGO	31	3.039	404	58,8	283	41,2	687	
3	KAB. TRENGGALEK	22	2.172	129	54,0	110	46,0	239	
4	KAB. TULUNGAGUNG	32	4.396	413	57,7	303	42,3	716	
5	KAB. BLITAR	24	4.917	276	56,9	209	43,1	485	
6	KAB. KEDIRI	37	4.038	675	53,5	587	46,5	1.262	
7	KAB. MALANG	39	6.384	990	56,2	772	43,8	1.762	
8	KAB. LUMAJANG	25	7.184	634	51,9	588	48,1	1.222	
9	KAB. JEMBER	50	15.900	1.602	52,9	1.426	47,1	3.028	
10	KAB. BANYUWANGI	45	5.715	1.007	53,2	885	46,8	1.892	
11	KAB. BONDOWOSO	25	5.250	480	51,4	454	48,6	934	
12	KAB. SITUBONDO	20	6.736	478	52,3	436	47,7	914	
13	KAB. PROBOLINGGO	33	3.027	619	51,9	574	48,1	1.193	
14	KAB. PASURUAN	33	12.734	904	52,9	804	47,1	1.708	
15	KAB. SIDOARJO	27	11.702	1.594	58,8	1.119	41,2	2.713	
16	KAB. MOJOKERTO	27	11.717	626	59,8	420	40,2	1.046	
17	KAB. JOMBANG	34	7.681	702	56,4	542	43,6	1.244	

Gambar 3 Jumlah Kasus Penyakit dalam Bentuk *PDF*

Penerapan Metode *K-Means*

K-Means adalah metode pengelompokan yang termasuk dalam bidang *unsupervised machine learning*. Teknik *unsupervised machine learning* mengidentifikasi pola dalam data tanpa bergantung pada label atau output yang telah ditentukan sebelumnya. *K-Means* bekerja dengan mengelompokkan data ke dalam beberapa *cluster* berdasarkan kemiripannya, sehingga data dalam satu *cluster* memiliki lebih banyak kesamaan karakteristik dibandingkan dengan data dalam *cluster* lainnya. Berikut langkah kinerja algoritma *K-Means*, disajikan pada bagan air Gambar 4.



Gambar 4 Bagan Air Kinerja *K-Means Cluster*ing

Nilai awal untuk mengelompokkan data point ke dalam 3 *cluster* berdasarkan pengetahuan domain untuk mencapai tujuan analisis. Merepresentasikan tingkat penyebaran penyakit berdasarkan daerah, yaitu daerah tinggi, sedang, dan rendah. Selanjutnya melakukan perhitungan jarak antara setiap titik dengan *centroid* menggunakan metode jarak Euclidean, rumusnya pada persamaan (1)

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \qquad (1)$$

dengan d(p,q) adalah Euclidean distance antara dua titik p dan q, p adalah data point pertama, q adalah data point kedua, n adalah jumlah dimensi atau atribut dalam data. Contoh perhitungan untuk objek data ke-1 pada setiap *cluster* adalah sebagai berikut:

$$Cluster0 = \sqrt{(214-214)^2 + (611-611)^2 + (362-362)^2 + (288-288)^2 + (212-212)^2} = 0$$

$$Cluster1 = \sqrt{(214-977)^2 + (611-1065)^2 + (362-1158)^2 + (288-904)^2 + (212-687)^2} = 1423{,}721181$$

$$Cluster2 = \sqrt{(214-401)^2 + (611-465)^2 + (362-552)^2 + (288-402)^2 + (212-239)^2} = 325{,}7452993$$

Setiap daerah dicari nilai terkecil di antara 3 *cluster* yang ada. Nilai yang ditampilkan dengan latar belakang warna hijau mencerminkan angka terkecil dan merepresentasikan jarak terdekat dari objek data ke *centroid* (pusat *cluster*), dapat dilihat pada Tabel 2.

Tabel 2 Hasil Perhitungan Jarak Data ke Centroid pada Iterasi Pertama

No	Nama Daerah	*Cluster*0	*Cluster*1	*Cluster*2
1	KAB. PACITAN	0,00	1423,7	325,75
2	KAB. PONOROGO	1423,72	0,0	1229,52
3	KAB. TRENGGALEK	325,75	1229,5	0,00
4	KAB. TULUNGAGUNG	1523,86	182,4	1343,82
5	KAB. BLITAR	713,29	777,9	514,64
…	…	…	…	…
34	KAB. LAMONGAN	3731,86	2335,1	3545,93
35	KAB. GRESIK	3858,37	2476,8	3689,22
36	KAB. SUMENEP	2751,31	1441,8	2592,58
37	KOTA MALANG	3134,67	1730,0	2946,75
38	KOTA SURABAYA	13055,28	11673,7	12888,79

Selanjutnya hitung ulang *centroid* di setiap *cluster*, hingga mencapai konvergensi. Konvergensi adalah situasi di mana lokasi *centroid* tidak banyak bergerak di antara iterasi, menunjukkan bahwa algoritma telah sampai pada solusi yang stabil. Nilai 0,0001 adalah ambang batas perubahan posisi centroid yang digunakan untuk memutuskan kapan iterasi algoritma *K-Means* harus dihentikan. Dihasilkan dari penyesuaian berdasarkan kebutuhan khusus penelitian

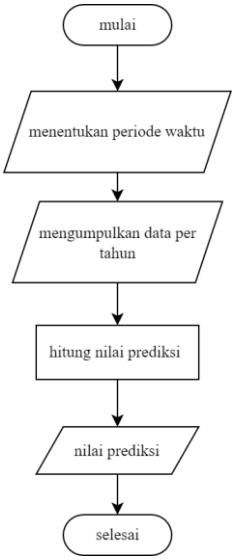
Cluster berisi nilai 0, 1, 2 berasal dari penempatan posisi *cluster*. Selanjutnya agar lebih jelas keterangan pengelompokannya, menggunakan fungsi statistik minimal dan maksimal. Nilai minimal masuk ke dalam *cluster* rendah berwarna hijau. Nilai maksimal akan masuk ke dalam *cluster* tinggi berwarna merah. Nilai berada di antara keduanya akan termasuk ke dalam *cluster* sedang berwarna kuning. Hasil akhir proses direpresentasikan dalam Tabel 3 setelah menjalani 14 kali iterasi.

Tabel 3 Hasil Akhir Proses *Clustering*

No	Nama Daerah	2017	2018	2019	2020	2021	*Cluster*	Keterangan
1	KOTA BATU	32	200	241	185	140	0	rendah
2	KAB. PROBOLINGGO	1371	999	1682	1165	1193	0	rendah
3	KAB. PONOROGO	977	1065	1158	904	687	0	rendah
4	KAB. TULUNGAGUNG	1043	1188	1236	821	716	0	rendah
5	KAB. NGAWI	835	811	1046	755	660	0	rendah
…	…	…	…	…	…	…	…	…
24	KOTA SURABAYA	6338	7007	7950	4151	4631	1	tinggi
…	…	…	…	…	…	…	…	…
34	KAB. PASURUAN	2393	2735	3181	1770	1708	2	sedang
35	KOTA MALANG	1783	1818	2218	1377	1342	2	sedang
36	KAB. JOMBANG	1089	1522	1677	1288	1244	2	sedang
37	KAB. GRESIK	2115	2305	2505	1463	1771	2	sedang
38	KAB. KEDIRI	1735	1709	1874	1542	1262	2	sedang

Penerapan Metode *Simple Moving Average*

Berikut langkah-langkah dalam urutan kinerja algoritma *Simple Moving Average*, disajikan pada bagan air Gambar 5. Jumlah periode waktu sebelumnya yang akan digunakan adalah 5.



Gambar 5 Bagan Air Kinerja *Simple Moving Average*

Selanjutnya menghitung rata-rata dari nilai-nilai dalam periode yang ditentukan dengan rumus persamaan (2).

$$P = \frac{x_1+x_2+x_3+x_4+\ldots+x_n}{n} \quad (2)$$

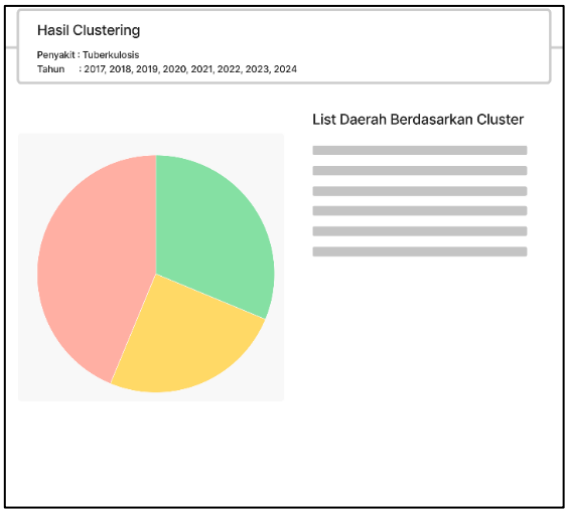
dengan P adalah prediksi *Simple Moving Average*, … adalah nilai dalam deret waktu untuk periode yang ditentukan, n adalah jumlah periode. Hasil akhir proses bisa dilihat pada Tabel 4.

Tabel 4 Hasil Akhir Prediksi *Simple Moving Average*

Nama Daerah	2017	…	2022	2023	2024
KAB. PACITAN	214	…	337	362	312
KAB. PONOROGO	977	…	958	954	932
KAB. TRENGGALEK	401	…	412	414	404
KAB. TULUNGAGUNG	1043	…	1001	992	953
KAB. BLITAR	462	…	627	660	653
…	…	…	…	…	…
KOTA PASURUAN	499	…	551	561	560
KOTA MOJOKERTO	497	…	431	418	423
KOTA MADIUN	634	…	602	595	576
KOTA SURABAYA	6338	…	6015	5951	5740
KOTA BATU	32	…	160	185	182

Visualisasi Data

Visualisasi menjadi alat untuk menyampaikan hasil penelitian yang rumit kepada pihak Dinas Kesehatan Provinsi Jawa Timur sebagai pemangku kepentingan. Menggambarkan data dengan cara yang mudah dipahami, seperti diagram lingkaran untuk menunjukkan pengelompokan pada Gambar 6, dan diagram garis untuk menunjukkan hasil prediksi pada Gambar 7.

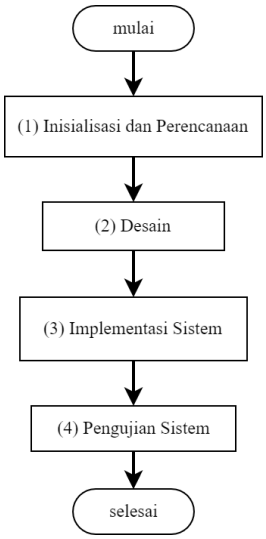


Gambar 6 Diagram Lingkaran Hasil *Clustering* Metode *K-Means*

Gambar 7 Diagram Garis Hasil Prediksi Metode *Simple Moving Average*

2.3 Tahap Pengembangan Sistem

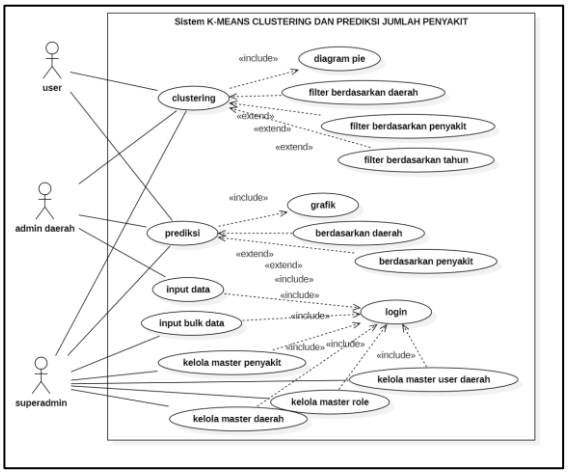
Proses pengembangan sistem dibagi menjadi 4 tahap yang dipresentasikan dalam bentuk bagan air pada Gambar 8.



Gambar 8 Bagan Air Tahap Pengembangan Sistem

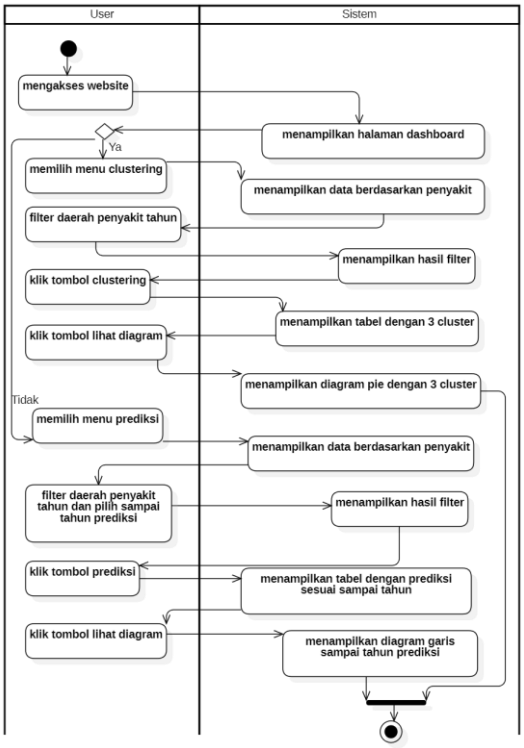
Pengembangan sistem ini dimulai dengan tahap inisialisasi dan perencanaan, yang meliputi penilaian kebutuhan, penetapan tujuan, dan perancangan arsitektur sistem.

Tahap kedua adalah proses desain sistem menggunakan UML (*Unified Modeling Language*), meliputi pembuatan *use case diagram* pada Gambar 9. Melibatkan 3 pengguna yaitu *user*, admin daerah dan superadmin.

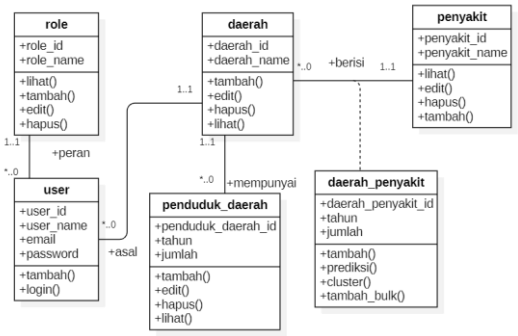


Gambar 9 *Use Case Diagram*

Selanjutnya terdapat *activity diagram user* pada Gambar 10, dan *class diagram* pada Gambar 11.



Gambar 10 *Activity Diagram User*



Gambar 11 *Class Diagram* Sistem

Tahap implementasi dimulai dengan pembagian kerja berdasarkan teknologi yang digunakan. *Frontend* dibangun dengan *Angular*, sebuah *framework JavaScript* yang kuat untuk membuat aplikasi web dinamis. *Backend* dibangun dengan *PHP (Hypertext Preprocessor)*, bahasa pemrograman pengembangan web sisi *server* karena fleksibilitas dan kemudahannya untuk dipelajari. Untuk manajemen data, *MySQL* dipilih karena merupakan sistem manajemen basis data relasional yang dapat diandalkan dan efisien. Sistem kemudian tersedia melalui penyedia *hosting* seperti *Hostinger*.

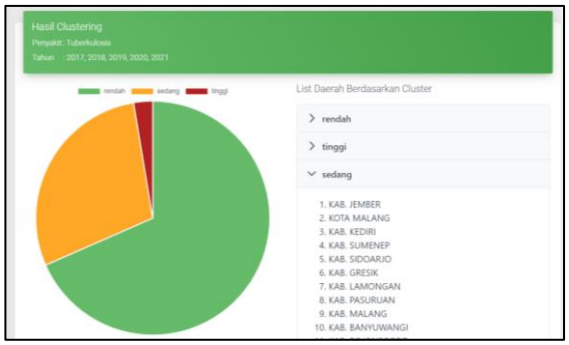
Tahap terakhir adalah pengujian sistem, yang terdiri dari pengujian fungsional menggunakan metode *blackbox*, pengujian kompatibilitas antar-*browser* menjamin bahwa program berfungsi dengan baik di berbagai *browser*. *Feedback* dari pengguna juga dikumpulkan untuk memastikan bahwa sistem memenuhi kebutuhan.

3. HASIL DAN PEMBAHASAN

Fokus utama dari bab hasil dan pembahasan adalah hasil dari penggunaan metode *K-Means Cluster*ing dan *Simple Moving Average* untuk menganalisis dan memprediksi jenis penyakit di Provinsi Jawa Timur. Sistem dibuat hanya untuk membantu implementasi kedua metode tersebut.

Metode *K-Means Cluster*ing berhasil digunakan untuk menggabungkan data penyakit dari 38 kota/kabupaten di Jawa Timur ke dalam 3 *cluster*: tinggi, sedang, dan rendah. Data penyakit dianalisis dengan menggunakan metode *K-Means* dalam proses *iteratif* yang melibatkan penghitungan dan pembaruan *centroid* untuk setiap *cluster* hingga konvergensi tercapai.

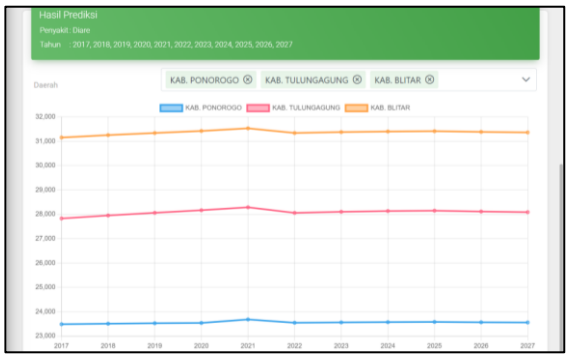
3 *cluster* yang berbeda menandakan area dengan tingkat keparahan penyakit yang tinggi, sedang, dan rendah. *Cluster* tinggi berisi lokasi dengan insiden penyakit sangat tinggi, *cluster* sedang berisi area dengan insiden penyakit sedang, dan *cluster* rendah berisi area dengan insiden penyakit relatif rendah. Hasil visualisasi data *clustering* dapat dilihat pada Gambar 12.



Gambar 12 Visualisasi Data Hasil *Clustering*

Analisis *cluster* menunjukkan bahwa *cluster* tinggi membutuhkan penanganan kesehatan segera, *cluster* sedang membutuhkan pemantauan berkelanjutan dan upaya pencegahan, dan *cluster* rendah memprioritaskan kegiatan pencegahan untuk menjaga kondisi tetap stabil.

Prediksi tren menggunakan metode *Simple Moving Average* menunjukkan bahwa jumlah kasus penyakit di setiap lokasi akan meningkat, menurun, atau tetap stabil, sangat berguna untuk perencanaan strategis dalam kegiatan pencegahan dan pengobatan penyakit di masa depan. Hasil visualisasi data prediksi dapat dilihat pada Gambar 13.



Gambar 13 Visualisasi Data Hasil Prediksi

Hasil penelitian menunjukkan bahwa metode yang digunakan memberikan kontribusi signifikan terhadap analisis dan manajemen penyakit di Provinsi Jawa Timur. Hasil pengelompokan menunjukkan bahwa pendekatan *K-Means* berhasil mengelompokkan wilayah berdasarkan tingkat keparahan penyakit, dengan jumlah *cluster* (k=3), sehingga memberikan hasil yang bermakna dan dapat diimplementasikan. Prediksi *Simple Moving Average* memberikan wawasan yang berguna tentang tren penyakit di masa depan, yang dapat membantu pemerintah daerah dan organisasi kesehatan untuk merancang strategi pencegahan dan pengobatan yang lebih efektif.

Pengujian

Hasil pengujian fungsional dilihat pada Tabel 5.

Hasil pengujian kompatibilitas aplikasi pada *Windows* 10 dengan 4 varian *browser* yang berbeda: *Microsoft Edge* versi 118, *Mozilla Firefox* versi 118,

Google Chrome versi 118, *Opera* versi 101, *Safari* versi 17. Selanjutnya hasil perbandingan semua *browser* ditunjukkan pada Tabel 6.

Tabel 5 Hasil Pengujian Fungsi *User*

No	Skenario	Hasil yang Diharapkan	Hasil Pengujian
1.	Memilih menu *dashboard*	Sistem menampilkan halaman *home*, grafik dan tabel	Berhasil
2.	Memilih menu *clustering*	Sistem menampilkan halaman *clustering* dan tabel	Berhasil
3.	Proses *clustering*	Sistem menampilkan hasil *clustering* penyakit tuberkulosis	Berhasil
4.	Lihat diagram *clustering*	Sistem menampilkan diagram lingkaran sesuai dengan hasil *clustering*	Berhasil
5.	Memilih menu prediksi	Sistem menampilkan halaman prediksi dan tabel	Berhasil
6.	Proses prediksi salah	Muncul pesan peringatan "Silahkan pilih tahun mulai minimal 3 tahun terakhir"	Berhasil
7.	Proses prediksi benar	Sistem menampilkan hasil prediksi penyakit diare hingga tahun 2026	Berhasil
8.	Lihat diagram garis	Sistem menampilkan diagram garis sesuai dengan hasil prediksi	Berhasil

Tabel 6 Hasil Uji Kompatibilitas Antar *Browser*

No	Nama *Browser*	Kompatibilitas	Fungsional	Catatan
1	*Microsoft Edge*	Kompatibel	*Cluster*ing, Prediksi, Visualisasi	Tidak ada masalah
2	*Mozilla Firefox*	Kompatibel	*Cluster*ing, Prediksi, Visualisasi	Tidak ada masalah
3	*Google Chrome*	Kompatibel	*Cluster*ing, Prediksi, Visualisasi	Tidak ada masalah
4	*Opera*	Kompatibel	*Cluster*ing, Prediksi, Visualisasi	Tidak ada masalah
5	*Safari*	Kompatibel	*Cluster*ing, Prediksi, Visualisasi	Masalah kecil dengan *rendering* grafis

Tabel 7 Rangkuman Hasil Evaluasi Kuisioner

Aspek Uji	Nilai	
P1	4,34 (Sangat Baik)	4,34 (Sangat Baik)
P2	4,28 (Sangat Baik)	
P3	4,36 (Sangat Baik)	
P4	4,44 (Sangat Baik)	
P5	4,36 (Sangat Baik)	
P6	4,36 (Sangat Baik)	
P7	4,3 (Sangat Baik)	

Pengujian pengguna melibatkan 50 responden, dibagi menjadi beberapa kelompok. Kelompok tenaga kesehatan seperti dokter dan perawat, pegawai Dinas Kesehatan yang bertanggung jawab atas administrasi dan manajemen data, peneliti kesehatan seperti akademisi dan mahasiswa, dan pengembang teknologi informasi yang berhubungan dengan kesehatan seperti *programmer* dan analis sistem. Tabel 7 merangkum hasil evaluasi kuesioner berdasarkan aspek *usability framework WebQual* 4.0.

Pertanyaan aspek uji P1 (Pertanyaan 1) adalah apakah pengguna merasa mudah untuk memahami cara mengoperasikan situs web, P2 adalah apakah interaksi antara situs web dan pengguna jelas dan sederhana, P3 adalah apakah mudah bagi pengguna untuk menjelajahi situs web, P4 adalah apakah pengguna merasa situs web mudah digunakan, P5 adalah apakah *website* memiliki tampilan yang menarik, P6 adalah apakah desainnya sesuai dengan jenis situs web, P7 adalah apakah pengguna merasa mudah untuk menemukan informasi yang mereka.

4. KESIMPULAN DAN SARAN

Penelitian ini berhasil menerapkan *K-Means Cluster*ing membantu dalam segmentasi data kasus penyakit di Provinsi Jawa Timur ke dalam 3 *cluster* yaitu rendah, sedang dan tinggi. Berhasil menerapkan metode *Simple Moving Average* untuk memprediksi tren kasus penyakit di Provinsi Jawa Timur untuk masa depan berdasarkan data historis dengan periode 5 tahun. Berhasil mengembangkan sistem berbasis *website* sebagai media implementasi prediksi dan *clustering* kasus penyakit di Provinsi Jawa Timur. *Website* mempunyai fitur penyaringan data berdasarkan penyakit, daerah, tahun dan menampilkan diagram lingkaran yang untuk hasil *clustering* dan diagram garis yang digunakan untuk hasil prediksi. Sistem berhasil dibuat dengan *database MySQL*, *backend PHP* dan *frontend Angular*.

Berdasarkan hasil pengujian kompatibilitas antar *browser* adalah sistem dapat diakses di berbagai jenis *browser* pada *Windows* 10 seperti *Microsoft Edge* versi 118, *Mozilla Firefox* versi 118, *Google Chrome* versi 118, *Opera* versi 101 dan *Safari* versi 17. Sedangkan hasil pengujian pengguna menggunakan metode *usability* dengan survei berbasis *WebQual* 4.0, yang menghasilkan nilai rata-rata 4,34 (sangat baik). Secara keseluruhan sistem mendapatkan penilaian yang baik dari para pengguna, yang dibuktikan dengan nilai yang tinggi pada dimensi *usability*.

Penelitian di masa depan diharapkan berfokus pada pembaruan aplikasi secara teratur untuk mempertahankan relevansi dalam lingkungan komputasi yang berubah dengan cepat. Selain itu, melibatkan pengguna secara aktif dalam proses pengembangan membantu meningkatkan kualitas aplikasi secara signifikan. Pengembang juga harus mendiversifikasi sumber data dan melakukan evaluasi perkembangan teknologi secara berkala untuk menjamin bahwa aplikasi tetap relevan dan berjalan dengan baik di lingkungan yang terus berubah.

DAFTAR PUSTAKA

ADIPUTRA, I.N.M., 2022. CLUSTERING PENYAKIT DBD PADA RUMAH SAKIT DHARMA KERTI MENGGUNAKAN ALGORITMA K-MEANS. *INSERT: Information System and Emerging Technology Journal*, 2(2), p.99. https://doi.org/10.23887/insert.v2i2.41673.

AGOESTINA, F.S., TAMBUNAN, H.S. dan NASUTION, R.A., 2022. Pemanfaatan Algoritma *K-Means* untuk Pengelompokan Pasien Penyakit Infeksi Saluran Pernafasan Akut (ISPA). *SATESI: Jurnal Sains Teknologi dan Sistem Informasi*, 2(1), pp.60–69. https://doi.org/10.54259/satesi.v2i1.804.

AJI, B.G., SONDAWA, D.C.A., ANINDIKA, F.A. dan JANUARITA, D., 2022. Analisis Peramalan Obat Menggunakan Metode *Simple Moving Average*, Weighted Moving Average, Dan Exponential Smoothing. *JURIKOM (Jurnal Riset Komputer)*, 9(4), pp.959–965. https://doi.org/10.30865/jurikom.v9i4.4454.

ALI, A. dan MASYFUFAH, L., 2021. *CLUSTER*ISASI PASIEN BPJS DENGAN METODE K-MEANS CLUSTERING GUNA MENUNJANG PROGRAM JAMINAN KESEHATAN NASIONAL DI RUMAH SAKIT ANWAR MEDIKA BALONG BENDO SIDOARJO. *Jurnal Wiyata*, 8(1), pp.8–22.

ANDI, A., JULIANDY, C. dan DAVID, D., 2023. *Cluster*ing Analysis of Tweets About COVID-19 Using the *K-Means* Algorithm. *Sinkron*, 8(1), pp.543–533. https://doi.org/10.33395/sinkron.v8i1.12145.

ARYADI, A.R., ANDREAS, F. dan SARI, B.N., 2022. Analisis *Cluster K-Means* pada Data Rata-Rata Konsumsi Kalori dan Protein Menurut Provinsi dengan Metode Davies Bouldin Index. *Jurnal Pendidikan dan Konseling (JPDK)*, [*online*] 4(4), pp.5652–5661. https://doi.org/10.31004/jpdk.v4i4.6375.

AYU RAMADHANTY, D., RENITA SYAFITRI, ERRISSYA RASYWIR dan DESPITA Meisak, 2022. Implementasi Data Mining Untuk Menentukan Persediaan Stok Obat Di Apotek K-24 Menggunakan Metode *K-Means Cluster*ing. *Jurnal Informatika Dan Rekayasa Komputer(JAKAKOM)*, 2(1), pp.155–160. https://doi.org/10.33998/jakakom.2022.2.1.31.

FIRDAUS, M.I., AISYIYAH, P. dan DEVI, R., 2022. Implementation of *K-Means* Algorithm for Diseases *Cluster*ing in Elderly Posyandu Participants. *ATCSJ Applied Technology and Computing Science Journal*, [*online*] 5(1), pp.11–20. https://doi.org/10.33086/atcsj.X.

JULIANA, E., ALEYDA, V.N. dan YULIANA, Y., 2021. Penerapan Metode *Cluster*ing *K-Means* Untuk Membantu Menentukan Tingkatan Status Daerah Dampak *Covid-19*. *Jurnal MediaTIK*, 4(3), p.112. https://doi.org/10.26858/jmtik.v4i3.23698.

NABYLA, F. dan MUKRODIN, 2021. Data Mining Pengelompokan Biaya Pengeluaran Rumah Sakit Oleh BPJS Kesehatan Menggunakan Algoritma *K-Means*. *Smart Comp: Jurnalnya Orang Pintar Komputer*, 10(2), pp.79–83. https://doi.org/10.30591/smartcomp.v10i2.2128.

NUGROHO, M.R., HENDRAWAN, I.E. dan PUWANTORO, 2022. Penerapan Algoritma *K-Means* Untuk *Cluster*isasi Data Obat Pada Rumah Sakit ASRI. *JURNAL NUANSA INFORMATIKA*, 16(1), pp.125–133.

NURHIDAYAH, A., LESTARI, W. dan PURWANTO, E., 2022. Implementasi Data Mining Algoritma *K-Means* untuk *Cluster*ing Penyakit di RS Panti Waluyo Surakarta. *DutaCom*, 15(2). https://doi.org/10.47701/dutacom.v15i2.2009.

PURBA, N., PONINGSIH dan TAMBUNAN, H.S., 2021. Penerapan Algoritma *K-Means Cluster*ing Pada Penyebaran Penyakit Infeksi Saluran Pernapasan Akut (ISPA) di Provinsi Riau. *Journal of Information System Research (JOSH)*, 2(3), pp.220–226.

PURWANTO, B., NILOGIRI, A. dan WARDOYO, A.E., 2022. Penerapan Algoritma *K-Means Cluster*ing Untuk Pengelompokan Penyebaran Penyakit Tbc (Studi Kasus: Puskesmas Di Kabupaten Jember). *Jurnal Smart Teknologi*, 3(3), pp.273–285.

RANI, R.A. dan NOVIANTI, L., 2022. Sistem Informasi Geografis Pemetaan dan Pengelompokan Balita Berpotensi Gizi Buruk Tertinggi pada Wilayah Kota Prabumulih Menggunakan Metode *K-Means Cluster*ing (Studi Kasus: Dinas Kesehatan Kota Prabumulih). *InfoTekJar: Jurnal Nasional Informatikadan Teknologi Jaringan*, 6(2), pp.248–253.

AL RIVAN, M.E. dan SONARU, R.A., 2022. Perbandingan Metode *K-Means* dan GA *K-Means* untuk *Cluster*ing Dataset Heart Disease Patients. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 9(3), pp.2585–2597. https://doi.org/10.35957/jatisi.v9i3.2799.

ROKHIMAH, W. dan KISMIANTINI, 2022. ANALISIS CLUSTERING TINGKAT KERAWANAN WILAYAH TERHADAP KASUS PENYAKIT DI KABUPATEN SLEMAN DENGAN METODE K-MEANS.

Jurnal Kajian dan Terapan Matematika, 8(1), pp.25–37.

SARI, T.P., HANANTO, A.L., NOVALIA, E., TUKINO, T. dan HILABI, S.S., 2023. Implementasi Algoritma *K-Means* dalam Analisis *Cluster*isasi Penyebaran Penyakit Hiv/Aids. *Infotek : Jurnal Informatika dan Teknologi*, 6(1), pp.104–114. https://doi.org/10.29408/jit.v6i1.7423.

SEMBIRING, E.M., 2022. Penerapan *K-Means Cluster*ing Untuk Pengelompokan Penyebaran Demam Berdarah Dengue (DBD) Di Kabupaten Deli Serdang. *TIN: Terapan Informatika Nusantara*, 2(11), pp.673–677. https://doi.org/10.47065/tin.v2i11.1503.

SULISTIYO HIDAYAT, F., BERLIANA, R., AFFANDI, P., ZULIANA, V. dan PADILAH, T.N., 2022. Penerapan *K-Means Cluster*ing dalam Pengelompokan Kasus Tuberkulosis di Provinsi Jawa Barat. *Jurnal Ilmiah Wahana Pendidikan*, [*online*] 8(15), pp.213–227. https://doi.org/10.5281/zenodo.7049113.

SYAMFITHRIANI, T.S., MIRANTIKA, N. dan TRISUDARMO, R., 2023. Perbandingan Algoritma *K-Means* dan K-Medoids Untuk Pemetaan Daerah Penanganan Diare Pada Balita di Kabupaten Kuningan. *JURNAL SISTEM INFORMASI BISNIS*, 12(2), pp.132–139. https://doi.org/10.21456/vol12iss2pp132-139.

TAMBUNAN, T.M., ANITA, MIGI SULISTIONO dan ERRISSYA RASYWIR, 2022. Penerapan Data Mining Menggunakan Metode *K-Means* Untuk Menentukan Status *Covid-19* di Jambi. *Jurnal Manajemen Teknologi Dan Sistem Informasi (JMS)*, 2(2), pp.196–204. https://doi.org/10.33998/jms.2022.2.2.36.

TAMPUBOLON, S.S. dan HANDOKO, K., 2021. Penerapan Data Mining Analisa Penyakit Menular Pada Manusia. *Jurnal Comasie*, 5(7), pp.41–50.

TANTY, SERASI GINTING, B. dan SIMANJUNTAK, M., 2021. Pengelompokan Penyakit Pada Pasien Berdasarkan Usia Dengan Metode *K-Means Cluster*ing (Studi Kasus : Puskesmas Bahorok). *ALGORITMA: Jurnal Ilmu Komputer dan Informatika*, 5(2), pp.88–99. https://doi.org/http://dx.doi.org/10.30829/algoritma.v5i2.10508.

UTARI, D.T., 2021. Analisis Karakteristik Wilayah Transmisi *Covid-19* dengan Menggunakan Metode *K-Means Cluster*ing. *Jurnal Media Teknik dan Sistem Industri*, 5(1), p.25. https://doi.org/10.35194/jmtsi.v5i1.1220.