

Penerapan Data mining untuk Clustering Kondisi Sosial Ekonomi Berdasarkan Kepemilikan Jaminan Kesehatan Menggunakan Algoritma K-Means

Rika Sari^{1*}, Muhammad Yasin²

^{1,2} Teknik, Informatika, Universitas Asahan

*Penulis Korespondensi: saririka702@gmail.com¹

Abstract. This study discusses the application of the K-Means Clustering algorithm in analyzing the socioeconomic conditions of communities in North Sumatra Province based on health insurance coverage. The background of this study stems from the continuing gap in access to and coverage of health insurance, which is influenced by differences in socioeconomic conditions between regions. The purpose of this study is to identify community groups with different patterns of health insurance ownership, analyze their influence on socioeconomic conditions, and explain the application of the K-Means algorithm as a data clustering method. The data used was obtained from the North Sumatra Provincial Statistics Agency (BPS) for the 2021–2023 period, with variables including BPJS Health Insurance Premium Assistance Recipients, BPJS Health Insurance Non-Premium Assistance Recipients, Private Insurance, and Insurance from Companies/Offices. The analysis process was carried out through the stages of variable selection, initial centroid determination, distance calculation using Euclidean Distance, and iteration. The results of the study show that there are two main clusters, namely clusters with high health insurance ownership rates and clusters with low ownership rates. In terms of the BPJS Health Insurance Variable for Contribution Assistance Recipients, the cities of Medan, Samosir, Tanjungbalai, Asahan, Batu Bara, Binjai, and Toba are included in the high cluster, while Deli Serdang is in the low cluster. These findings are expected to provide an overview for local governments in formulating policies to improve access to health insurance in a more targeted and data-driven manner.

Keywords: BPJS; Health Insurance; K-Means Algorithm; North Sumatra Province; Socioeconomic Conditions.

Abstrak. Penelitian ini membahas penerapan algoritma *K-Means Clustering* dalam menganalisis kondisi sosial ekonomi masyarakat di Provinsi Sumatera Utara berdasarkan kepemilikan jaminan kesehatan. Latar belakang penelitian ini berangkat dari masih adanya kesenjangan akses dan cakupan jaminan kesehatan yang dipengaruhi oleh perbedaan kondisi sosial ekonomi antar daerah. Tujuan dari penelitian ini adalah untuk mengidentifikasi kelompok masyarakat dengan pola kepemilikan jaminan kesehatan yang berbeda, menganalisis pengaruhnya terhadap kondisi sosial ekonomi, serta menjelaskan penerapan algoritma *K-Means* sebagai metode pengelompokan data. Data yang digunakan diperoleh dari Badan Pusat Statistik (BPS) Provinsi Sumatera Utara periode 2021–2023, dengan variabel meliputi BPJS Kesehatan Penerima Bantuan Iuran, BPJS Kesehatan Non-Penerima Bantuan Iuran, Asuransi Swasta, dan Jaminan dari Perusahaan/Kantor. Proses analisis dilakukan melalui tahapan pemilihan variabel, penentuan centroid awal, perhitungan jarak menggunakan *Euclidean Distance*, serta iterasi. Hasil penelitian menunjukkan bahwa terdapat dua klaster utama, yaitu klaster dengan tingkat kepemilikan jaminan kesehatan tinggi dan klaster dengan tingkat kepemilikan rendah. Pada Variabel BPJS Kesehatan Penerima Bantuan Iuran Kota Medan, Samosir, Tanjungbalai, Asahan, Batu Bara, Binjai, dan Toba termasuk dalam klaster tinggi, sedangkan Deli Serdang berada pada klaster rendah. Temuan ini diharapkan dapat memberikan gambaran bagi pemerintah daerah dalam merumuskan kebijakan peningkatan akses jaminan kesehatan secara lebih tepat sasaran dan berbasis data.

Kata kunci: Algoritma *K-Means*; Asuransi Kesehatan; BPJS; Kondisi Sosial Ekonomi; Provinsi Sumatera Utara.

1. LATAR BELAKANG

Status sosial ekonomi suatu masyarakat sangat mempengaruhi aksesibilitas dan kualitas pelayanan kesehatan. Di Indonesia khususnya di Sumatera Utara jaminan kesehatan merupakan inisiatif utama pemerintah untuk meningkatkan kesejahteraan warganya. Meskipun program jaminan kesehatan telah dilaksanakan masih terdapat kesenjangan dalam cakupan

jaminan kesehatan di berbagai tingkat masyarakat. Hal ini menunjukkan bahwa faktor sosio-ekonomi berperan dalam kemampuan masyarakat untuk memperoleh jaminan kesehatan sehingga memerlukan analisis lebih lanjut untuk mengembangkan kebijakan yang lebih terfokus.

Masalah lain yang perlu diidentifikasi adalah bagaimana hasil clustering dapat digunakan untuk merumuskan kebijakan yang lebih efektif. Setelah kelompok-kelompok masyarakat teridentifikasi langkah selanjutnya adalah menganalisis karakteristik masing-masing kelompok dan merumuskan intervensi yang sesuai. Misalnya kelompok masyarakat dengan kondisi sosial ekonomi rendah yang memiliki akses terbatas terhadap jaminan kesehatan mungkin memerlukan program edukasi dan peningkatan kesadaran tentang pentingnya jaminan kesehatan. Oleh karena itu hasil penelitian ini diharapkan dapat memberikan rekomendasi yang baik untuk mengambil keputusan bijak.

2. KAJIAN TEORITIS

Kepemilikan Jaminan Kesehatan

Kepemilikan jaminan kesehatan merupakan indikator penting dalam menilai tingkat kesejahteraan masyarakat. Menurut *Undang-Undang Nomor 36 Tahun 2009 tentang Kesehatan*, setiap warga negara berhak memperoleh pelayanan kesehatan yang layak tanpa diskriminasi. Dalam konteks Indonesia, program Jaminan Kesehatan Nasional (JKN) yang diselenggarakan oleh BPJS Kesehatan menjadi salah satu upaya utama pemerintah untuk meningkatkan akses terhadap layanan kesehatan yang merata di seluruh wilayah.

Faktor-faktor yang memengaruhi kepemilikan jaminan kesehatan meliputi tingkat pendapatan, pendidikan, akses terhadap fasilitas kesehatan, serta kebijakan daerah. Menurut Jannah (2023), keberhasilan program jaminan kesehatan tidak hanya ditentukan oleh aspek kebijakan, tetapi juga oleh kesiapan infrastruktur kesehatan dan tingkat kesadaran masyarakat terhadap pentingnya perlindungan kesehatan.

Dengan demikian, analisis terhadap pola kepemilikan jaminan kesehatan dapat memberikan gambaran mengenai kesenjangan sosial ekonomi antarwilayah, serta menjadi dasar perumusan kebijakan publik yang lebih efektif.

Data Mining

Data mining merupakan proses penemuan pola tersembunyi, hubungan, atau informasi bermakna dari kumpulan data besar. Yanti, Yetri, dan Taufik (2024) mendefinisikan data mining sebagai penerapan metode statistik, algoritma pembelajaran mesin, dan teknik analisis data untuk menemukan pengetahuan baru yang tidak tampak secara eksplisit dalam data.

Proses data mining terdiri dari beberapa tahapan utama yang dikenal dengan Knowledge Discovery in Databases (KDD), yaitu:

- Pemilihan data, yakni pemilihan atribut yang relevan.
- Praproses dan transformasi data, untuk menghapus data tidak lengkap dan menyesuaikan format.
- Penerapan algoritma, misalnya klasifikasi atau clustering.
- Evaluasi pola, guna menilai validitas hasil.
- Visualisasi hasil, agar informasi mudah dipahami.

Clustering

Clustering merupakan teknik pengelompokan data yang bertujuan untuk menempatkan objek dengan karakteristik serupa ke dalam satu kelompok. Sitohang (2025) menjelaskan bahwa clustering berperan penting dalam menemukan struktur alami dalam data tanpa memerlukan label kelas awal.

Putri, Cobantoro, dan Mustikasari (2024) menyatakan bahwa metode clustering memungkinkan analisis segmentasi sosial berdasarkan kemiripan karakteristik ekonomi dan demografi. Hasil pengelompokan tersebut dapat dimanfaatkan untuk mengidentifikasi daerah prioritas pembangunan atau kelompok masyarakat yang perlu intervensi kebijakan khusus.

Algoritma K-Means

Algoritma K-Means merupakan salah satu metode clustering yang paling populer karena kemampuannya dalam mengelompokkan data numerik dengan efisien. Menurut Sitohang (2025), algoritma ini bekerja dengan cara menghitung jarak antar data terhadap titik pusat (centroid) dari masing-masing kelompok, kemudian memperbarui posisi centroid hingga hasilnya stabil. Langkah-langkah utama dalam algoritma K-Means meliputi:

Unified Modeling Language (UML)

Unified Modeling Language (UML) merupakan bahasa pemodelan standar yang digunakan untuk merancang dan menggambarkan sistem perangkat lunak secara visual. Karim Syahputra dan Asri Awaliyah Arifin (2024) menjelaskan bahwa UML membantu pengembang memahami hubungan antar komponen sistem, mulai dari peran pengguna, proses, hingga aliran data.

Berbagai diagram UML, seperti use case diagram, class diagram, activity diagram, dan sequence diagram, digunakan untuk menggambarkan interaksi antar elemen sistem. Ramadhani (2025) menambahkan bahwa UML berfungsi sebagai alat komunikasi yang efektif antara perancangan sistem dan pengguna, sehingga membantu memastikan sistem informasi yang dikembangkan sesuai dengan kebutuhan pengguna.

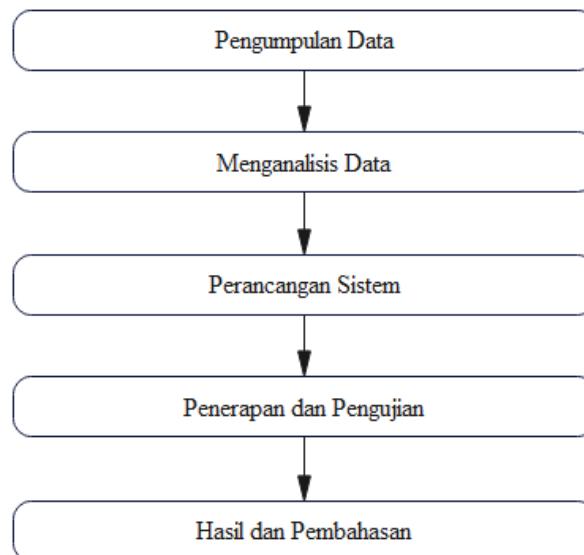
3. METODE PENELITIAN

Metode Dan Tempat Penelitian.

Metode ini menggunakan metode deskriptif sebagai teknik analisis data untuk menyelesaikan penelitian ini. Dengan mengungkapkan keadaan subjek atau objek dalam bentuk jurnal penelitian orang lain, kita dapat memecahkan kesulitan dengan menggunakan metode ini. Tempat penelitian adalah lokasi tempat penelitian dilakukan, tempat penelitian memiliki pengaruh yang signifikan terhadap cara penelitian dilakukan karena menyediakan bahan-bahan yang dibutuhkan untuk penelitian. Dalam hal ini, peneliti berfokus pada keadaan sosial ekonomi Sumatera Utara sebagaimana yang dipersepsikan oleh kepemilikan asuransi kesehatan. Adapun tempat yang akan diteliti merupakan Kota Medan, Deli Serdang, Samosir, Tanjung Balai, Asahan dan Batu Bara.

Kerangka Penelitian

Kerangka kerja ini memberikan panduan sistematis dalam melaksanakan penelitian mengenai penerapan data mining dalam clustering kondisi sosial ekonomi. Berikut ini kerangka kerja penelitian ini bisa dilihat pada gambar 1. dibawah ini:



Gambar 1. Kerangka Kerja Penelitian

Uraian Kerangka Kerja Penelitian

Pengumpulan Data

Pada tahapan ini, data kepemilikan jaminan kesehatan di kumpulkan dari sumber resmi Badan Pusat Statistik Provinsi Sumatra Utara (BPS SUMUT), melalui situs resmi mereka di (<https://sumut.bps.go.id/id>). Data yang diambil mencakup priode pada tahun 2021, 2022 dan 2023.

Menganalisis Data

Pada tahap ini akan dilakukan analisis terhadap permasalahan yang ada. Penelitian melakukan analisis data metode Algoritma K-Means untuk mendapatkan hasil Clustering kondisi sosial ekonomi di sumatera utara, dalam tahapan analisis data ini peneliti melakukan beberapa tahapan, diantaranya merupakan:

- Menentukan jumlah klaster (K).
- Menetapkan centroid awal secara acak.
- Menghitung jarak setiap data ke centroid (menggunakan rumus Euclidean Distance).
- Mengelompokkan data ke dalam klaster dengan jarak terdekat.
- Memperbarui posisi centroid dan mengulang proses sampai konvergen.

Perancangan Sistem

Perancangan sistem menggunakan PHP dan UML dimulai dengan analisis kebutuhan pengguna untuk memahami fungsionalitas yang diinginkan, di mana diagram *use case* UML digunakan untuk menggambarkan interaksi antara pengguna dan sistem. Setelah itu, dilakukan perancangan sistem dengan membuat diagram kelas yang mendefinisikan struktur data dan hubungan antar objek, diikuti dengan diagram urutan untuk merinci alur proses dalam sistem. Selanjutnya, pengembangan dilakukan dengan menggunakan PHP sebagai bahasa pemrograman untuk mengimplementasikan antarmuka pengguna, serta mengintegrasikan basis data untuk penyimpanan data.

Penerapan dan Pengujian

Data yang telah diolah menggunakan bahasa pemrograman PHP kini diimplementasikan dan diuji. Hal ini bertujuan untuk menjamin bahwa model ini akan bermanfaat bagi konsumen, sehingga metode ini dapat digunakan untuk mengetahui kondisi sosial ekonomi klaster di Sumatera Utara.

Hasil dan Pembahasan

Setelah sistem selesai dibuat, temuan dari prosedur pengujian data akan dijelaskan. Tahap pengujian dilakukan untuk memastikan bahwa setiap fitur berfungsi sebagaimana mestinya dan sesuai dengan persyaratan sebelum sistem akhirnya tersedia bagi pengguna. Pendekatan Algoritma K-Means telah digunakan untuk menguji data. Hasil pengelompokan akan menghasilkan tingkat persentase (%) akurasi dalam mengidentifikasi kondisi sosial yang tepat yang akan memengaruhi peningkatan standar hidup yang disediakan pemerintah.

4. HASIL DAN PEMBAHASAN

Selanjutnya merupakan langkah-langkah dalam proses k-means clustering. Dalam tahap ini yang dihitung pertama dari BPJS Kesehatan Penerima Bantuan Iuran dan dilanjutkan dengan data yang lainnya.

Perhitungan BPJS Penerima Bantuan Iuaran

Perhitungan dilakukan dengan menggunakan rumus Euclidean Distance $[(x,y),(a,b)] = \sqrt{(x - a)^2 + (y - b)^2}$.

Tabel 1. Tabel Data.

Kabupaten	BPJS Penerima Bantuan Iuaran		
	2021	2022	2023
Medan	30,03	28,7	39,01
Deli Serdang	21,65	38,83	25,31
Samosir	53,71	27,31	53,99
Tanjungbalai	57,03	34,15	57,13
Asahan	30,52	21,7	31,06
Batu Bara	39,59	19,55	34,9
Binjai	36,23	29,18	28,73
Toba	40,46	25,26	38,16

Iterasi pertama

Pilih centroid awal sebagai pusat cluster. Misalnya kita ambil dua centroid awal berikut ini;

Tabel 2. Iterasi pertama.

Nama	2021	2022	2023
C1	30.03	28.7	39.01
C2	21.65	38.83	25.31

Tabel 3. Menghitung Jarak Terhadap Pusat Cluster

Nama	C1	C2
K01	0	18.99
K02	18.19	0
K03	28.05	44.53
K04	32.97	47.81
K05	10.6	20.13
K06	13.86	28.03
K07	12.01	17.82
K08	11.02	26.52

Pada tahap ini menghitung jarak antar cluster dengan cara data yang ada pada C1 dan C2 dikurang dan ditambahkan secara bergantian dengan semua data yang ada, selanjutnya data yang udah dikurangkan dikuadratkan sehingga mendapatkan hasil yang ada pada tabel diatas.

Tabel 4. Keanggotaan Cluster dan Jarak Minimum

Nama	Keanggotaan	Minimum	Kuadrat
K01	C1	0	0
K02	C2	0	0
K03	C3	28.05	787.07
K04	C4	33.05	1092.45
K05	C5	10.6	112.44
K06	C6	13.86	192.01
K07	C7	12.01	144.35
K08	C8	11.02	121.34
	WCV		2449.66

Dalam menentukan keanggotaan dilihat dari nilai klaster jika tinggi maka termasuk dalam C1 dan rendah termasuk C2, untuk menentukan nilai minimum dilakukan pembagian nilai C1 dan C2. Selanjutnya nilai dari minimum dikuadratkan sehingga dapat dilihat hasilnya pada tabel diatas.

Tabel 5. Pusat Cluster Baru.

Nama	C1			C2		
	2021	2022	2023	2021	2022	2023
K01	30.03	28.7	39.01			
K02				21.65	38.83	25.31
K03	53.71	27.31	53.99			
K04	57.03	34.15	57.13			
K05	30.52	21.7	31.06			
K06	39.59	19.55	34.9			
K07	36.23	29.18	28.73			
K08	40.46	25.26	38.16			
Pusat Cluster	41.08	26.55	40.43	21.65	38.83	25.31

Dalam menentukan pusat cluster baru dilakukan pemisahan data antara C1 dan C2 selanjutnya dicari nilai rata-rata dari C1 dan C2 untuk menentukan pusat klaster.

Tabel 6. Jarak Antar Cluster Baru.

Cluster 1	Cluster 2	D
C1	C2	18.99
BVC		18.99
Ratio		0.0077

Cara menentukan jarak antar klaster baru dengan cara klaster sebelumnya C1 dan C2 dikurang dan ditambahkan secara bergantian kemudian dikuadratkan. Karena rasio (0.0077682901255939) > rasio sebelumnya (0), maka iterasi dilanjutkan.

Iterasi kedua

Tabel 7. Pusat Cluster.

Nama	2021	2022	2023
C1	41.08	26.55	40.43
C2	21.65	38.83	25.31

Dalam memusatkan klaster baru diambil dari nilai rata-rata klaster sebelumnya.

Tabel 8. Menghitung Jarak Terhadap Pusat Cluster.

Nama	C1	C2
K01	11.35	18.99
K02	27.51	0
K03	18.55	44.53
K04	24.31	47.81
K05	14.93	20.13
K06	9.04	28.03
K07	12.93	17.82
K08	2.68	26.52

Pada tahap ini menghitung jarak antar cluster dengan cara data yang ada pada C1 dan C2 dikurang dan ditambahkan secara bergantian dengan semua data yang ada, selanjutnya data yang udah dikurangkan dikuadratkan sehingga mendapatkan hasil yang ada pada tabel diatas.

Tabel 9. Keanggotaan Cluster dan Jarak Minimum.

Nama	Keanggotaan	Minimum	Kuadrat
K01	C1	11.35	128.76
K02	C2	0	0
K03	C1	18.55	344.05
K04	C1	24.31	591.15
K05	C1	14.93	222.78
K06	C1	9.04	81.76
K07	C1	12.93	167.24
K08	C1	2.68	7.18
	WCV		1542.93

Dalam menentukan keanggotaan dilihat dari nilai klaster jika tinggi maka termasuk dalam C1 dan rendah termasuk C2, untuk menentukan nilai minimum dilakukan pembagian nilai C1 dan C2. Selanjutnya nilai dari minimum dikuadratkan sehingga dapat dilihat hasilnya pada tabel diatas.

Tabel 10. Pusat Cluster Baru.

Nama	C1			C2		
	2021	2022	2023	2021	2022	2023
K01	30.03	28.7	39.01			
K02				21.65	38.83	25.31
K03	53.71	27.31	53.99			
K04	57.03	34.15	57.13			
K05	30.52	21.7	31.06			
K06	39.59	19.55	34.9			
K07	36.23	29.18	28.73			
K08	40.46	25.26	38.16			
Pusat Cluster	41.08	26.55	40.43	21.65	38.83	25.31

Dalam menentukan pusat cluster baru dilakukan pemisahan data antara C1 dan C2 selanjutkan dicari nilai rata-rata dari C1 dan C2 untuk menentukan pusat klaster.

Tabel 11. Jarak Antar Cluster Baru.

Cluster 1	Cluster 2		D
	C1	C2	
		27.51	27.51
BVC			0.0178
Ratio			

Cara menentukan jarak antar klaster baru dengan cara klaster sebelumnya C1 dan C2 dikurang dan ditambahkan secara bergantian kemudian dikuadratkan. Karena rasio (0.01783050062992) > rasio sebelumnya (0.0077682901255939), maka iterasi dilanjutkan.

Iterasi ketiga

Tabel 12. Pusat Cluster Iterasi Ketiga

Nama	2021	2022
C1	41.08	26.55
C2	21.65	38.83

Dalam memusatkan klaster baru diambil dari nilai rata-rata klaster sebelumnya.

Tabel 13. Menghitung Jarak Terhadap Pusat Cluster

Nama	C1	C2
K01	11.35	18.99
K02	27.51	0
K03	18.55	44.53
K04	24.31	47.81
K05	14.93	20.13
K06	9.04	28.03
K07	12.93	17.82
K08	2.68	26.52

Pada tahap ini menghitung jarak antar cluster dengan cara data yang ada pada C1 dan C2 dikurang dan ditambahkan secara bergantian dengan semua data yang ada, selanjutnya data yang udah dikurangkan dikuadratkan sehingga mendapatkan hasil yang ada pada tabel diatas.

Tabel 14. Keanggotaan Cluster dan Jarak Minimum

Nama	Keanggotaan	Minimum	Kuadrat
K01	C1	11.35	128.76
K02	C2	0	0
K03	C1	18.55	344.05
K04	C1	24.31	591.15
K05	C1	14.93	222.78
K06	C1	9.04	81.76
K07	C1	12.93	167.24
K08	C1	2.68	7.18
	WCV		1542.93

Dalam menentukan keanggotaan dilihat dari nilai klaster jika tinggi maka termasuk dalam C1 dan rendah termasuk C2, untuk menentukan nilai minimum dilakukan pembagian nilai C1 dan C2. Selanjutnya nilai dari minimum dikuadratkan sehingga dapat dilihat hasilnya pada tabel diatas.

Tabel 15. Pusat Cluster Baru.

Nama	C1			C2		
	2021	2022	2023	2021	2022	2023
K01	30.03	28.7	39.01			
K02				21.65	38.83	25.31
K03	53.71	27.31	53.99			
K04	57.03	34.15	57.13			
K05	30.52	21.7	31.06			
K06	39.59	19.55	34.9			
K07	36.23	29.18	28.73			
K08	40.46	25.26	38.16			
Pusat Cluster	41.08	26.55	40.43	21.65	38.83	25.31

Dalam menentukan pusat cluster baru dilakukan pemisahan data antara C1 dan C2 selanjutkan dicari nilai rata-rata dari C1 dan C2 untuk menentukan pusat klaster.

Tabel 16. Jarak Antar Cluster Baru

Cluster 1	Cluster 2	D
C1	C2	27.51
BVC		27.51
Ratio		0.017

Karena rasio (0.01783050062992) Tidak > rasio sebelumnya (0.01783050062992), maka iterasi dihentikan. Maka didapatkan hasil dari Iterasi ke: 3 dengan Rasio : 0.01783050062992

Tabel 17. Hasil Dari Cluster BPJS Kesehatan Penerima Bantuan Iuran

Kode	Nama	Centroid	Keterangan
K01	Kota Medan	C1	Tinggi
K02	Deli Serdang	C2	Rendah
K03	Samosir	C1	Tinggi
K04	Tanjungbalai	C1	Tinggi
K05	Asahan	C1	Tinggi
K06	Batu Bara	C1	Tinggi
K07	Binjai	C1	Tinggi
K08	Toba	C1	Tinggi

- Adapun jumlah cluster C1 adalah 7 bahan pokok yaitu Kota Medan, Samosir, Tanjungbalai, Asahan, Batu Bara, Binjai, Toba
- Adapun jumlah cluster C2 adalah 1 bahan pokok yaitu Deli Serdang

5. KESIMPULAN

Penelitian ini menunjukkan bahwa penerapan algoritma K-Means Clustering efektif dalam menganalisis kondisi sosial ekonomi masyarakat Provinsi Sumatera Utara berdasarkan kepemilikan jaminan kesehatan. Berdasarkan hasil pengelompokan, diperoleh dua klaster utama, yaitu klaster dengan tingkat kepemilikan jaminan kesehatan tinggi dan klaster dengan tingkat kepemilikan rendah. Daerah yang termasuk dalam klaster tinggi meliputi Kota Medan, Samosir, Tanjungbalai, Asahan, Batu Bara, Binjai, dan Toba, sedangkan Deli Serdang termasuk dalam klaster rendah. Temuan ini menggambarkan adanya kesenjangan sosial ekonomi antarwilayah yang berpengaruh terhadap kepemilikan jaminan kesehatan. Hasil penelitian diharapkan dapat menjadi dasar bagi pemerintah daerah dalam merumuskan kebijakan yang lebih tepat sasaran untuk meningkatkan pemerataan akses dan kualitas pelayanan jaminan kesehatan bagi seluruh masyarakat Sumatera Utara.

DAFTAR REFERENSI

- Amalia, I., Ffitria, Chalvina, & Sitorus, M. (2025). *Analisa dan pengelompokan kota/kabupaten di Indonesia berdasarkan aspek sosial ekonomi menggunakan algoritma K-Means*. Innovation and Technology. (Artikel dalam persiapan)
- Indaswari, R. A., Akbar, Z., & Santoso, H. (2024). *Sistem aplikasi e-Booking berbasis web pada Misumi Beauty & Skin*. Jurnal Informatika, Sistem Informasi dan Kehutanan (FORSINTA), 3(2).
- Munawarah, R., Anggraini, W. A., Azzahra, D., & Pramita, F. (2023). *Faktor yang mempengaruhi pemanfaatan Jaminan Kesehatan Nasional pada layanan kesehatan Puskesmas (literature review)*. Jurnal Kesehatan Tambusai, 4(3), 2229–2237.

- Pinaria, A. (2022). *Sistem informasi pengarsipan surat berbasis website*. JUSTIFY Jurnal ..., ...
- Pratama, C. A. (2017). *Dasar pemrograman web dengan PHP*. Jakarta: Deepublish.
- Putri, I., Fajaryanto C., & Mustikasari, D. (n.d.). *Penggunaan algoritma K-Means untuk menentukan calon penerima beasiswa*. (Belum diterbitkan formal)
- Ramdani, R., Suarna, N., Ali, I., & Efendi, D. I. (2025). *Penerapan algoritma K-Means dalam analisis data kependudukan untuk optimalisasi pengelompokan di Desa Pasawahan*. Jurnal Informatika dan Teknik Elektro Terapkan, 13(1). <https://doi.org/10.23960/jitet.v13i1.5639>
- Selvi, D., & Hidayanti, I. (2024). *Perancangan sistem informasi buku tamu pada instansi Badan Penelitian dan Pengembangan Daerah Provinsi Sumatera Selatan*. MDP Student Conference, 3(1), 88–94.
- Sianturi, M. W. (2025). *Penerapan analisis cluster K-Means untuk pengelompokan kabupaten/kota di Provinsi Sumatera Utara berdasarkan indikator kemiskinan tahun 2023*. Jurnal ..., Vol ..., No ..., ...
- Sinlae, F., Maulana, I., Setiyansyah, F., & Ihsan, M. (2024). *Pengenalan pemrograman web: Pembuatan aplikasi web sederhana dengan PHP dan MySQL*. Jurnal Siber Multi Disiplin (JSMD), 2(2), 68–82. <https://doi.org/10.38035/jsmd.v2i2.156>
- Sitohang, P. (2025). *Klasterisasi peserta Olimpiade Lomba Kompetensi Siswa (LKS) menggunakan algoritma K-Means*. Interdisiplin Journal Social Science (IJSS), 1, 33–40. <https://doi.org/10.33330/ijss.v1i1.7>
- Syahputra, K., Abdul, & Awaliyah Arifin, A. (2024). *Sistem informasi pengarsipan surat berbasis web pada media dan informasi Universitas Royal*. JAMUDI Jurnal Multidisiplin Ilmu, ... (...), ... (halaman belum disebut)
- Wahyudi, G. R., Rahmaddeni, E., & Dini, E. (2025). *Pengelompokan kabupaten/kota di Indonesia untuk pemetaan kondisi ekonomi daerah berdasarkan PDRB menggunakan algoritma K-Means*. MALCOM-Jurnal ..., 5(3), 1143-1151. <https://doi.org/10.57152/malcom.v5i3.2206> [Jurnal IRPI](#)
- Yanti, E., Milfa Y., & Taufik, F. (2024). *Penerapan data mining untuk mengestimasi ...* Jurnal Cyber Tech, 8(...), 1103–1108. <https://doi.org/10.36040/jati.v8i1.8214>
- Zahra, A. F., Annisa, & Suwanda, R. (2023). *Aplikasi buku tamu berbasis website di PT PLN (Persero) ULP Lhoksukon*. Jurnal Rekayasa Sistem Informasi dan Teknologi, 1(2), 188–195. <https://doi.org/10.59407/jrsit.v1i2.218> [Journal PPMI](#)