Jurnal Matematika UNAND
Vol. **14** No. **1** Hal. 93 – 102
Edisi Januari 2025
ISSN : 2303–291X
e-ISSN : 2721–9410

https://doi.org/10.25077/jmua.14.1.93-102.2025

CLUSTERING ANALYSIS OF PROVINCIAL IN INDONESIA BASED ON THE 2023 HUMAN DEVELOPMENT INDEX INDICATORS USING THE K-MEDOIDS ALGORITHM

SYAFA MARWA SABRIN, TABAH HERI SETIAWAN*
Program Studi Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Pamulang
email : syafamarwa13@gmail.com, tabah.ibnubara@gmail.com

Diterima 26 Agustus 2024 Direvisi 23 Desember 2024 Dipublikasikan 31 Januari 2025

Abstrak. Indonesia memiliki visi Indonesia Emas pada tahun 2045, namun pencapaian Indeks Pembangunan Manusia (IPM) dalam 20 tahun terakhir menunjukkan tantangan untuk mewujudkan visi tersebut. Penelitian ini menggunakan algoritma *k-medoids* untuk melakukan *clustering* provinsi di Indonesia berdasarkan indikator IPM tahun 2023. *K-medoids* dipilih karena keunggulannya dalam menangani *outlier*. Berdasarkan hasil perbandingan dengan metode *k-means* dan *fuzzy c-means*, metode *k-medoids* terbukti merupakan metode terbaik, karena *cluster* yang terbentuk pada *k-medoids* terpisah dengan baik dan memiliki truktur yang kuat. Hasil penelitian menghasilkan tiga *cluster*: C_1 memiliki anggota provinsi dengan IPM sangat tinggi, C_2 memiliki anggota provinsi dengan IPM tinggi, sementara C_3 memiliki anggota provinsi dengan IPM sedang. Analisis ini diharapkan menjadi bahan evaluasi dan referensi bagi pengembangan metode *clustering*.

Kata Kunci: Clustering, Indeks Pembangunan Manusia, K-medoids

Abstract. *Indonesia has a vision of Golden Indonesia 2045, yet the achievement of the Human Development Index (HDI) over the past 20 years indicates challenges in realizing this vision. This study employs the k-medoids algorithm to cluster provinces in Indonesia based on HDI indicators for 2023. K-medoids was chosen for its advantages in handling non-normally distributed data and being more robust to outliers. A comparison with k-means and fuzzy c-means methods confirmed that k-medoids is the best approach, as the clusters formed are well-separated and demonstrate strong structure. The results identified three clusters: C_1 consists of provinces with very high HDI, C_2 includes provinces with high HDI, while C_3 comprises provinces with moderate HDI. This analysis is expected to serve as an evaluation and reference for further clustering method.*

Keywords: Clustering, Human Development Index, K-medoids

*Penulis Korespondensi

1. Pendahuluan

Indonesia memiliki visi untuk menjadi negara yang tangguh, makmur, inklusif, dan berkelanjutan. Untuk mewujudkan visi tersebut, KADIN Indonesia telah merumuskan Peta Jalan (*Road Map*) Indonesia Emas 2045 [1]. Kementerian PPN/Bappenas juga telah mengembangkan Rencana Pembangunan Jangka Panjang Nasional (RPJPN) 2025-2045 dalam upaya mewujudkan visi tersebut [2].

Indeks Pembangunan Manusia (IPM) dapat menjadi salah satu acuan dalam keberhasilan visi tersebut, karena sasaran pokok dari pembangunan manusia adalah untuk mengupayakan peningkatan kesejahteraan manusia secara adil, inklusif, dan berkelanjutan. Dalam mengukur kualitas hidup, IPM didasari oleh tiga dimensi penilaian yang meliputi dimensi umur panjang dan hidup sehat, dimensi pengetahuan, dan dimensi kehidupan yang layak [3]. Pada laporan IPM terbaru yang diterbitkan oleh *United Nations Development Programme* (UNDP) pada tahun 2022, Indonesia meraih nilai IPM sebesar 0,713 yang termasuk dalam kategori negara dengan IPM yang tinggi [4]. Sayangnya, nilai IPM Indonesia tersebut ternyata di bawah rata-rata IPM global yakni 0,739, menjadikan Indonesia berada di peringkat 112 dari 193 negara. Selain itu, IPM Indonesia juga tertinggal dari beberapa negara ASEAN, di antaranya Singapura, Brunei Darussalam, Malaysia, Thailand, dan Vietnam [5].

Dalam mengatasi masalah ini, pemerintah telah melakukan pengelompokan untuk menempatkan daerah ke dalam suatu kelompok yang memiliki kemiripan dalam pencapaian pembangunan manusia berdasarkan IPM. Dimana hasil pengelompokan tersebut dapat digunakan sebagai acuan bagi pemangku kebijakan dalam merumuskan prioritas pembangunan daerah di setiap provinsi. Agar provinsi yang berada pada kategori IPM sedang dapat ditingkatkan menjadi kategori tinggi, provinsi dengan kategori IPM tinggi dapat ditingkatkan menjadi kategori sangat tinggi, dan mempertahankan provinsi yang sudah berada pada kategori IPM sangat tinggi.

Akan tetapi, pengelompokan (*clustering*) memiliki berbagai macam metode dalam menganalisis data dengan pendekatan dan kelebihannya sendiri. Oleh karena itu, dibutuhkan pembandingan metode yang berbeda untuk melihat konsistensi hasil *cluster* secara lebih lanjut [6]. Pada penelitian ini digunakan metode *k-medoids clustering*, metode ini dipilih karena keunggulannya dalam mengatasi beberapa keterbatasan metode *clustering* lainnya, seperti sensitivitas terhadap *outlier* dan kemampuannya dalam menangani data yang tidak berdistribusi normal [7]. Untuk evaluasi lebih lanjut dan juga melihat konsistensi hasil *cluster* maka akan dilakukan pula perbandingan menggunakan metode lain di antaranya yaitu metode *k-means* dan *fuzzy c-means*.

2. Landasan Teori

2.1. *Analisis Cluster*

Analisis *cluster* yaitu sebuah teknik analisis variabel berganda dengan sasaran utama melakukan pengelompokan beberapa objek yang memiliki karakteristik serupa ([8], [9]). Karakteristik yang dimiliki oleh setiap objek dalam sebuah *cluster* memiliki tingkat kemiripan yang tinggi, sedangkan karakteristik yang dimiliki oleh objek dalam *cluster* berbeda memiliki tingkat kemiripan yang rendah. Dapat di-

artikan bahwa perbedaan di dalam sebuah *cluster* adalah minimum sedangkan perbedaan antar *cluster* adalah maksimum [10]. Pada penelitian ini digunakan jarak *euclidean* yang merupakan pengukuran kedekatan dua buah objek data secara geometris. Jarak *euclidean* dihitung dengan menggunakan persamaan sebagai berikut ([11], [12]):

$$d(x,y) = \sqrt{\sum_{i=1}^{z}(x_{ik} - y_{jk})^2}, \tag{2.1}$$

dimana $d(x,y)$ adalah jarak antara x dan y, i adalah data individual, z adalah total data, x_{i_k} adalah pusat data *cluster*, dan y_{j_k} adalah data di setiap jk data.

Analisis *cluster* pada umumnya dibedakan menjadi dua, yaitu *hierarchical clustering* dan *partitional clustering*. *Hierarchical clustering* merupakan pengelompokan data berdasarkan bagan hirarki, dimana setiap iterasi atau pembagian seluruh *data set* dibagi menjadi beberapa *cluster* terjadi penggabungan dua kelompok terdekat. Sedangkan *partitional clustering* adalah metode dimana data dikelola menjadi sekumpulan *cluster* tanpa terdapat keterkaitan hirarki antara masing-masing *cluster* ([13], [14]).

2.2. *K-Medoids*

K-medoids clustering ialah metode pengklasteran parsial yang serupa dengan *k-means clustering*. *K-medoids clustering* memperkecil jarak antara kumpulan data pada sebuah *cluster* dengan kumpulan data yang representatif untuk *cluster* sejenis. Kumpulan titik data yang representatif ini disebut *medoid*. Dengan demikian, *k-medoids clustering* menjamin bahwa pusat dari sebuah *cluster* adalah titik data yang paling terpusat ([15], [16], [17]).

2.3. *K-Means*

K-means Clustering merupakan metode pengelompokan data non hirarki yang mengelompokkan data menjadi satu atau lebih *cluster*. Data yang memiliki karakteristik serupa dikelompokkan dalam sebuah *cluster* berbeda sehingga data dalam satu kelompok *cluster* memiliki tingkat kemiripan yang kecil. Algoritma *k-means Clustering* menggunakan proses iteratif untuk mendapatkan basis *cluster* ([18], [19]).

2.4. *Fuzzy C-Means*

Fuzzy c-means ialah metode yang digunakan untuk pengelompokan data yang bersifat *supervised*, dimana *cluster* ditentukan berdasarkan *input cluster* dan hasil *output* yang diharapkan. Derajat keanggotaan sebuah *cluster* menentukan apakah sebuah titik data termasuk dalam *fuzzy c-means*, sebuah metode pengelompokan data [20].

2.5. *Silhouette Coefficient*

Silhouette coefficient merupakan metode untuk mengetahui kekuatan dan kualitas *cluster*, seberapa baik sebuah objek berada dalam sebuah *cluster*. Penggunaan

metode ini merupakan kombinasi dari metode kohesi dan separasi. Nilai *silhouette coefficient* dari sebuah objek i adalah sebagai berikut [21]:

$$S_i = \frac{b_i - a_i}{\max\{a_i, b_i\}}, \tag{2.2}$$

dimana a_i adalah jarak rata-rata untuk objek i dengan seluruh objek pada *cluster* yang sama dan b_i adalah jarak rata-rata minimum untuk objek i dengan seluruh objek pada *cluster* terdekat.

Adapun kriteria-kriteria pengukuran nilai *silhouette coefficient* dibuat oleh Kauffman dan Rousseuw [22], dan diberikan pada Tabel 1 sebagai berikut.

Tabel 1. [22] Interpretasi *Silhouette Coefficient*

Nilai *Silhouette Coefficient*	**Interpretasi**
0.71 - 1.00	Stuktur Kuat
0.51 - 0.70	Struktur Baik
0.26 - 0.50	Struktur Lemah
0.00 - 0.25	Tidak Terstruktur

2.6. *Evaluasi Metode Clustering*

Evaluasi metode yang memiliki kinerja terbaik dilakukan dengan mempertimbangkan rasio rata-rata standar deviasi di dalam *cluster* dan standar deviasi antar *cluster*. Standar deviasi rata-rata dalam *cluster* (S_w) dinyatakan dengan:

$$S_w = \frac{1}{c}\sum_{k=1}^{c} S_k. \tag{2.3}$$

Standar deviasi antar *cluster* (S_w) dinyatakan dengan:

$$S_b = \left[\frac{1}{c-1}\sum_{k=1}^{c}\left(\overline{X}_k - \overline{X}\right)^2\right]^{\frac{1}{2}}, \tag{2.4}$$

dimana c adalah jumlah *cluster*, S_k adalah standar deviasi di dalam *cluster* ke-k, $\overline{X}_k$ adalah rata-rata *cluster* ke-k, dan $\overline{X}$ adalah rata-rata dari seluruh *cluster*. Semakin kecil nilai S_w dan semakin besar nilai S_b, maka metode tersebut memiliki kinerja yang baik, artinya memiliki homogenitas yang tinggi [23].

3. Pembahasan

3.1. *Penentuan Jumlah Cluster Optimal*

Dalam menetapkan jumlah *cluster* yang optimal, umumnya dihitung menggunakan *silhouette coefficient* untuk berbagai jumlah *cluster*, lalu memilih jumlah *cluster* yang menghasilkan nilai *silhouette coefficient* tertinggi. Pada Tabel 2 diberikan hasil perhitungan nilai *silhouette coefficient* dalam setiap metode.

Tabel 2. Hasil Nilai *Silhouette Coefficient*

Jumlah Cluster *k*	**Nilai *Silhouette Coefficient***		
	K-Medoids	***K-Means***	***Fuzzy C-Means***
2	0.4127	0.3073	0.3073
3	0.5398	0.3164	0.2877
4	0.4156	0.2888	0.2627
5	0.3352	0.2904	0.2572

Berdasarkan Tabel 2, diperoleh nilai tertinggi untuk k optimum pada metode *k-medoids* adalah $k = 3$ dengan stuktur *cluster* baik, untuk metode *k-means* adalah $k = 3$ dengan struktur *cluster* lemah, dan untuk metode *fuzzy c-means* adalah $k = 2$ dengan struktur *cluster* lemah.

3.2. *Penentuan dan Evaluasi Metode Clustering*

Dalam melakukan evaluasi metode *clustering* dapat ditentukan dari rata-rata standar deviasi di dalam *cluster* (S_w) dan standar deviasi antar *cluster* (S_b). Semakin rendah nilai rasio S_w terhadap S_b, berarti metode tersebut memiliki kinerja yang baik dan tingkat homogenitas yang tinggi di dalam cluster. Hasil perhitungan rasio S_w terhadap S_b dari setiap metode diberikan pada Tabel 3 berikut.

Tabel 3. Hasil Rasio rasio S_w terhadap S_b

Metode *Clustering*	**Jumlah** k	$\mathbf{S}_w/\mathbf{S}_b$
K-Medoid	$k = 3$	1.0875
K-Means	$k = 3$	0.5053
Fuzzy C-Means	$k = 2$	1.4626

Berdasarkan Tabel 3, diketahui bahwa nilai rasio yang paling rendah adalah metode *k-means*, yang menunjukkan bahwa *cluster* yang terbentuk lebih seragam dalam hal sebaran data. Selanjutnya, *k-medoids* memiliki rasio simpangan baku yang lebih tinggi dibandingkan dengan *k-means*, nilai ini menunjukkan bahwa *cluster* yang terbentuk cukup seragam, tetapi ada sedikit perbedaan ukuran antara *cluster*. Sedangkan pada metode *fuzzy c-means* memiliki nilai rasio simpangan baku yang paling tinggi, yang menunjukkan bahwa data dalam *cluster* lebih tersebar, sehingga *cluster* lebih kabur dan tidak terlalu terdefinisi dengan jelas.

3.3. *Penerapan Cluster K-Medoids*

Melalui *silhouette coefficients* dan rasio standar deviasi, dapat diketahui bahwa metode yang paling baik adalah *k-medoids clustering*. Karena, meskipun nilai rasio standar deviasi pada *k-means* paling rendah, namun nilai *silhouette coefficients* pada *k-medoids* lebih tinggi yang menunjukkan bahwa *cluster* yang terbentuk pada

k-medoids terpisah dengan baik. Nilai *silhouette coefficients* pada *cluster k-medoids* sebesar 0.5398 juga menunjukkan bahwa *cluster* tersebut memiliki struktur yang baik, sedangkan *k-means* memiliki struktur yang lemah.

Selain itu, nilai *silhouette coefficient* memberikan gambaran yang lebih baik mengenai performa *clustering* secara keseluruhan karena mencakup *separation* (jarak antar *cluster*) dan *cohesion* (kekompakan antar *cluster*). Sedangkan rasio standar deviasi hanya mengukur variabilitas dalam ukuran *cluster*, tanpa memperhitungkan jarak antar *cluster* [24]. Pada Tabel 4 diberikan hasil keanggotaan *cluster* yang diperoleh menggunakan metode *k-medoids*:

Tabel 4. Hasil Pembentukan *Cluster K-Medoids*

Provinsi	**Jarak Ke K_1**	**Jarak Ke K_2**	**Jarak Ke K_3**	**Jarak Terdekat**	***Cluster***
Aceh	2.7735	2.0330	1.6950	1.6950	C_1
Sumatera Utara	1.8976	1.9104	0.6897	0.6897	C_1
Sumatera Barat	2.2148	2.2277	1.1161	1.1161	C_1
Riau	1.2995	1.9896	0.0000	0.0000	C_1
Jambi	0.7894	1.7694	0.6526	0.6526	C_1
...	...	...	...	...	...
Sulawesi Barat	1.9768	1.1301	2.4399	1.1301	C_3
Maluku	3.4820	1.7705	2.6342	1.7705	C_3
Maluku Utara	2.7291	0.8830	2.2376	0.8830	C_3
Papua Barat	3.4323	1.7933	3.6371	1.7933	C_3
Papua	4.3692	4.0163	5.2628	4.0163	C_3

Dari hasil keanggotaan *cluster* di atas, diketahui *cluster* yang terbentuk adalah sebanyak 3 *cluster*, dimana terdapat 15 provinsi pada C_1, 9 provinsi pada C_2, dan 10 provinsi pada C_3. Untuk mengetahui lebih detail hasil *clustering* tersebut dapat dilihat melalui Tabel 5.

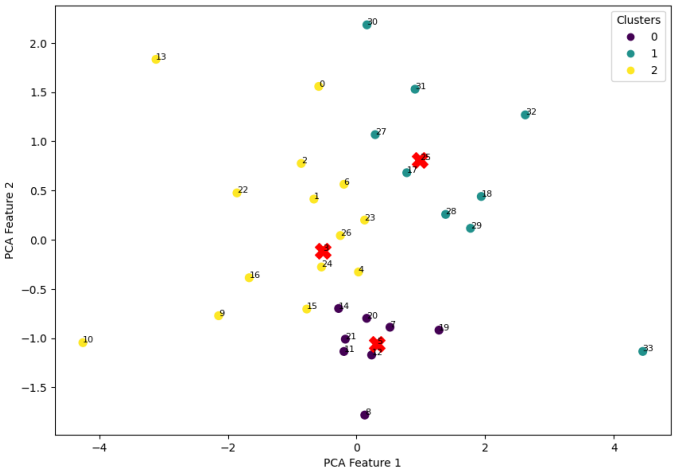
Selanjutnya, hasil pembentukan *cluster* tersebut divisualisasikan dengan menggunakan *cluster plot*. Dapat dilihat pada Gambar 1 dimana C_1 (*cluster* 2) ditandai dengan titik-titik berwarna kuning, C_2 (*cluster* 0) ditandai dengan titik-titik berwarna ungu, dan C_3 (*cluster* 1) ditandai dengan titik-titik berwarna hijau.

3.4. *Interpretasi Hasil Cluster*

Setiap *cluster* memiliki karakteristik yang berbeda berdasarkan kemiripan objeknya. Hal ini dapat ditentukan dengan menghitung rata-rata (*mean*) setiap variabel IPM pada setiap *cluster*. Mengacu pada ketentuan proses standarisasi *z-score*, data yang terlampir pada Tabel 6 menunjukkan bahwa angka negatif (-) menggambarkan nilai data di bawah keseluruhan rata-rata, dan angka positif (+) menggambarkan nilai data di atas keseluruhan rata-rata.

Tabel 5. Keanggotaan *Cluster*

Cluster	**Provinsi**	**Jumlah Anggota**
C_1	Aceh, Bali, Banten, Bengkulu, D.I. Yogyakarta, DKI Jakarta, Jambi, Kalimantan Timur, Kalimantan Utara, Kepulauan Riau, Riau, Sulawesi Selatan, Sulawesi Utara, Sumatera Barat, Sumatera Utara.	15
C_2	Jawa Barat, Jawa Tengah, Jawa Timur, Kalimantan Barat, Kalimantan Selatan, Kalimantan Tengah, Kepulauan Bangka Belitung, Lampung, Sumatera Selatan.	9
C_3	Gorontalo, Maluku, Maluku Utara, Nusa Tenggara Barat, Nusa Tenggara Timur, Papua, Papua Barat, Sulawesi Barat, Sulawesi Tengah, Sulawesi Tenggara.	10



Gambar 1. Visualisasi Hasil *Cluster*.

Tabel 6. Interpretasi Variabel IPM

	UHH	**HLS**	**RLS**	**PRP**
C_1	0.5774	0.4712	0.7454	0.5090
C_2	0.5676	-0.7263	-0.6640	0.1588
C_3	-1.3769	-0.0531	-0.5205	-0.9064

dimana:

UHH : Umur Harapan Hidup,
HLS : Harapan Lama Sekolah,
RLS : Rata-rata Lama Sekolah,
PRP : Pengeluaran Riil Per-Kapita.

Adapun karakteristik dari setiap *cluster* yang diperoleh adalah sebagai berikut:

(1) *Cluster* 1 : C_1 memiliki karakteristik nilai rata-rata seluruh variabel yaitu UHH, HLS, RLS, dan PRP berada di atas kesluruhan rata-rata provinsi. Karakteristik ini sesuai dengan provinsi-provinsi di C_1 yang memiliki nilai IPM sangat tinggi.
(2) *Cluster* 2 : Karakteristik C_2 memiliki nilai rata-rata variabel UHH dan PRP berada di atas keseluruhan rata-rata Provinsi, sedangkan variabel HLS dan RLS berada di bawah keseluruhan rata-rata Provinsi. Karakteristik ini sesuai dengan provinsi-provinsi di C_2 yang memiliki nilai IPM tinggi.
(3) *Cluster* 3 : Pada C_3, terbentuk karakteristik bahwa nilai rata-rata seluruh variabel yaitu UHH, HLS, RLS dan PRP berada di bawah keseluruhan rata-rata provinsi. Karakteristik ini sesuai dengan provinsi-provinsi di C_3 yang memiliki nilai IPM sedang.

4. Kesimpulan

Berdasarkan hasil nilai *silhouette coefficient* dan rasio standar deviasi diperoleh bahwa metode *clustering* terbaik adalah *k-medoids*. Karena, meskipun nilai rasio standar deviasi pada *k-means* paling rendah, namun nilai *silhouette coefficients* pada *k-medoids* lebih tinggi yang menunjukkan bahwa *cluster* yang terbentuk pada *k-medoids* terpisah dengan baik. Nilai *silhouette coefficients* pada *cluster k-medoids* sebesar 0.5398 juga menunjukkan bahwa *cluster* tersebut memiliki struktur yang baik, sedangkan *k-means* memiliki struktur yang lemah.

Adapun *cluster* yang terbentuk menggunakan metode *k-medoids* diperoleh sebanyak 3 *cluster*. Dimana terdapat 15 anggota *cluster* pada C_1 yang merupakan provinsi dengan IPM sangat tinggi. Pada C_2 terdapat 9 anggota *cluster* yang merupakan provinsi dengan IPM tinggi. Serta terdapat 10 anggota *cluster* pada C_3 yang merupakan provinsi dengan IPM sangat sedang. Melalui karakteristik setiap *cluster*, dapat diketahui variabel mana yang perlu menjadi prioritas pemerintah dalam upaya peningkatan dan pemerataan pembangunan.

5. Ucapan Terima kasih

Penulis berterima kasih kepada semua pihak yang telah membantu dalam penelitian ini, baik secara langsung maupun tidak langsung. Terutama kepada BPS (Badan Pusat Statistik) yang telah menyediakan data penelitian, serta kepada Universitas Pamulang yang telah membantu dan mendukung pada proses penelitian.

Daftar Pustaka

[1] KADIN Indonesia, 2024, Peta Jalan (*Road Map*) Indonesia Emas 2045, https://kadin.id/program/indonesia-emas/
[2] Kementerian PPN/Bappenas, 2024, Rancangan Akhir RPJPN 2025-2045 Indonesia Emas 2045, https://indonesia2045.go.id/
[3] BPS, 2015, Indeks Pembangunan Manusia, https://halbarkab.bps.go.id/subject/26/indeks-pembangunan-manusia.html
[4] UNDP, 2022, Human Development Index (HDI), https://hdr.undp.org/data-center/human-development-index#/indicies/HDI

[5] Ahdiat, A., 2024, Indeks Pembangunan Manusia ASEAN 2022, Indonesia Tak Menonjol, https://databoks.katadata.co.id/datapublish/2024/03/18/indeks-pembangunan-manusia-asean-2022-indonesia-tak-menonjol

[6] Sutramiani, N. P., Arthana, I. M. T., Lampung, P. F., Aurelia, S., Fauzi, M., Darma, I. W. A. S., 2024, The Performance Comparison of DBSCAN and K-Means Clustering for MSMEs Grouping based on Asset Value and Turnover, *Journal of Information Systems Engineering and Business Intelligence*, Vol. **10**(1): 13 – 24

[7] Asmiatun, S., Wakhidah, N., Putri, A. N., 2020, Penerapan Metode K-Medoids Untuk Pengelompokan Kondisi Jalan Di Kota Semarang, *Jurnal Teknik Informatika Dan Sistem Informasi*, Vol. **6**(2): 171 – 180

[8] Wijayanti, W., HG, I, R., Yanuar, F., 2021, Penggunaan Metode Fuzzy C-Means Untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Kesehatan Lingkungan, *Jurnal Matematika UNAND*, Vol. **10**(1): 129 – 136

[9] Ginantra, N. L. W. S. R., Arifah, F. N., Wijaya, A. H., Septarini, R. S., Ahmad, N., Ardiana, D. P. Y., Effendy, F., Iskandar, A., Hazriani, Sari, I. Y., Gustiana, Z., Prianto, C., Gustian, D., Negara, E. S., 2021, *Data Mining dan Penerapan Algoritma*, Yayasan Kita Menulis, Medan

[10] Kyrychenko, O., Ostapov, S., Malyk, I., 2023, Cluster Analysis of Information in Complex Networks. *International Journal of Computing*, Vol. **22**(4): 515 – 523

[11] Badung, N. M. A. A., Fernandes, A. A. R., Nugroho, W. H., 2021, Comparison of Distance and Linkage in Integrated Cluster Analysis with Multiple Discriminant Analysis on Home Ownership Credit Bank in Indonesia. *Mathematics and Statistics*, Vol. **9**(6): 958 – 975

[12] Dani, A. T. R., Wahyuningsih, S., Rizki, N. A., 2021, Grouping of Time Series Data using Cluster Analysis, *EKSPONENSIAL*, Vol. **11**(1): 29 – 38

[13] Irwansyah, E., Faisal, M., 2019, *Advanced Clustering : Teori dan Aplikasi*, Deepublish, Yogyakarta

[14] Amna, S, W., Sudipa, I. G. I., Putra, T. A. E., Wahidin, A. J., Syukrilla, W. A., Heryana, N., Indriyani, T., Santoso, L. W., 2023, *Data Mining*, PT. Global Eksekutif Teknologi, Padang

[15] Wanto, A., Siregar, M. N. H., Windarto, A. P., Hartama, D., Ginantra, N. L. W. S. R., Napitupulu, D., Negara, E. S., Lubis, M. R., Dewi, S. V., Prianto, C., 2020, *Data Mining : Algoritma dan Implementasi*, Yayasan Kita Menulis, Medan

[16] Prahara, A., Ismi, D. P., Azhari, A., 2020, Parallelization of Partitioning Around Medoids (PAM) in K-Medoids Clustering on GPU. *Knowledge Engineering and Data Science*, Vol. **3**(1): 40 – 49

[17] Luchia, N. T., Handayani, H., Hamdi, F. S., Erlangga, D., Octavia, S. F., 2022, Comparison of K-Means and K-Medoids on Poor Data Clustering in Indonesia. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, Vol. **2**(2): 35 – 41

[18] Pratiwi, Y., Mulyawan, M., 2023, Implementasi Algoritma K-Means untuk Menentukan Angka Harapan Hidup berdasarkan Tingkat Provinsi. *Blend Sains Jurnal Teknik*, Vol. **1**(4): 284 – 294

[19] Hashim, D. K., Muhammed, L. A. N., 2022, Performance of K-means algorithm based an ensemble learning. *Bulletin of Electrical Engineering and Informatics*,

Vol. **11**(1): 575 – 580

[20] Nasrullah, A. H., 2024, Klasterisasi Kabupaten Berdasarkan Index Pendidikan Penduduk Menggunakan Fuzzy C-Means. *Indonesian Technology and Education Journal*, Vol. **2**(2): 187 – 199

[21] Izzaty, U., Rahmi HG, I., Devianto, D., 2020, Pengklasteran Kabupaten/Kota di Provinsi Sumatera Barat Berdasarkan Indikator Kesejahteraan Masyarakat dengan Validitas Koefisien Silhouette, *Jurnal Matematika UNAND* Vol. **9**(2): 192 – 198

[22] Sirodj, D. A. N., Sumertajaya, I. M., Kurnia, A., 2023, Analisis Clustering Time Series untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indeks Pembangunan Manusia Jenis Kelamin Perempuan, *STATISTIKA Journal of Theoretical Statistics and Its Applications*, Vol. **23**(1): 29 – 37

[23] Ningrat, D. R., Maruddani, D. A. I., Wuryandari, T., 2016, Analisis Cluster dengan Algoritma K-Means dan Fuzzy C-Means Clustering untuk Pengelompokan Data Obligasi Korporasi, *Jurnal Gaussian*, Vol. **5**(4): 641 – 650

[24] Vardakas, G., Pavlopoulos, J., Likas, A., 2024, Revisiting Silhouette Aggregation, https://doi.org/10.48550/arXiv.2401.05831