

ANALISIS CLUSTER HIERARKI UNTUK PEMETAAN KESENJANGAN DIGITAL ANTAR PROVINSI DI INDONESIA

Andini Puti Aria

S-1 Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu, andini.puti@mhs.unib.ac.id

M. Syarlan

S-1 Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu, m.syarlan@mhs.unib.ac.id

Febi Herlizalia

S-1 Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu, febi.herlizalia@mhs.unib.ac.id

M. Zidan Mezilano Y.

S-1 Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu,
muhammadzidan141104@gmail.com

Novi Puspita*

S-1 Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Bengkulu, npuspita@unib.ac.id

Abstrak

Kesenjangan digital antarwilayah di Indonesia masih menjadi tantangan serius, terutama dalam hal akses terhadap telepon seluler, komputer, dan internet. Penelitian ini bertujuan memetakan kesenjangan digital antar provinsi di Indonesia berdasarkan tiga indikator akses teknologi informasi dan komunikasi menggunakan metode *hierarchical clustering*. Data yang digunakan bersumber dari Badan Pusat Statistik tahun 2024 yang mencakup 38 provinsi dengan variabel persentase pengguna telepon seluler, persentase pengguna komputer, dan persentase akses internet dalam tiga bulan terakhir. Metode yang digunakan adalah *hierarchical clustering* dengan pendekatan *agglomerative*, di mana lima metode *linkage* (*single*, *complete*, *average*, *Ward.D2*, dan *centroid*) dibandingkan menggunakan korelasi *cophenetic*. Hasil penelitian menunjukkan bahwa metode *centroid* menghasilkan korelasi *cophenetic* tertinggi sebesar 0,9341 dan terpilih sebagai metode terbaik. Berdasarkan *Elbow Method*, *Silhouette Method*, dan *NbClust*, jumlah *cluster* optimal yang terbentuk adalah dua *cluster*. *Cluster 1* beranggotakan 36 provinsi dengan rata-rata akses internet 89,73%, telepon seluler 69,08%, dan komputer 19,66%. *Cluster 2* hanya terdiri dari Papua Tengah dan Papua Pegunungan dengan nilai yang jauh lebih rendah pada seluruh indikator. Kualitas *cluster* dikonfirmasi oleh rata-rata *silhouette width* sebesar 0,79 untuk *Cluster 1* dan 0,61 untuk *Cluster 2*.

Kata Kunci: kesenjangan digital, *hierarchical clustering*, akses internet, pemetaan wilayah, Indonesia.

Abstract

Digital disparities between regions in Indonesia remain a serious challenge, particularly regarding access to mobile phones, computers, and the internet. This study aims to map the digital divide among provinces in Indonesia based on three indicators of information and communication technology access using the *hierarchical clustering* method. The data used were sourced from Statistics Indonesia in 2024, covering 38 provinces with variables including the percentage of mobile phone users, the percentage of computer users, and the percentage of internet access within the last three months. The method employed was *hierarchical clustering* with an *agglomerative* approach, where five *linkage* methods (*single*, *complete*, *average*, *Ward.D2*, and *centroid*) were compared using the *cophenetic* correlation coefficient. The results showed that the *centroid* method produced the highest *cophenetic* correlation of 0.9341 and was selected as the best method. Based on the *Elbow Method*, *Silhouette Method*, and *NbClust*, the optimal number of clusters formed was two. *Cluster 1* consisted of 36 provinces with average internet access of 89.73%, mobile phone ownership of 69.08%, and computer usage of 19.66%. *Cluster 2* consisted only of Central Papua and Mountain Papua, with much lower values across all indicators. Cluster quality was confirmed by an average *silhouette width* of 0.79 for *Cluster 1* and 0.61 for *Cluster 2*.

Keywords: digital divide, *hierarchical clustering*, internet access, regional mapping, Indonesia.

PENDAHULUAN

Kesenjangan digital telah menjadi isu krusial di Indonesia karena perbedaan akses terhadap teknologi informasi dan komunikasi (TIK) berdampak signifikan terhadap kesejahteraan sosial masyarakat (Jayanthi & Dinaseviani, 2022). Fenomena ini tidak hanya mencerminkan ketidaksetaraan dalam penguasaan teknologi, tetapi juga berpotensi memperburuk disparitas ekonomi antarwilayah. Menurut data APJII, jumlah pengguna internet di Indonesia pada tahun 2024 mencapai 221,56 juta jiwa dengan tingkat penetrasi 79,5% dari total populasi (Widiasanti et al., 2025).

Ketimpangan akses internet antarprovinsi di Indonesia masih menjadi tantangan serius, di mana penetrasi internet di daerah perkotaan mencapai 69,5% sementara di pedesaan hanya 30,5%. Provinsi di kawasan timur Indonesia seperti Papua memiliki tingkat akses internet yang jauh lebih rendah dibandingkan DKI Jakarta (Widiasanti et al., 2025). Kondisi geografis Indonesia yang terdiri atas berbagai pulau turut memperparah ketidakmerataan pembangunan infrastruktur TIK (Jayanthi & Dinaseviani, 2022).

Variabel yang digunakan meliputi persentase penduduk yang memiliki telepon seluler, persentase rumah tangga yang memiliki komputer, dan persentase rumah tangga yang mengakses internet per provinsi. Ketiga indikator ini merepresentasikan dimensi akses fisik dan material dasar dalam mengukur keterjangkauan masyarakat terhadap teknologi digital. Kepemilikan ponsel tidak selalu berarti kesetaraan dalam penggunaan digital, karena keterampilan yang dimiliki juga menentukan peluang individu.

Metode *analisis cluster* hierarki dipilih karena kemampuannya mengelompokkan provinsi tanpa penetapan jumlah klaster di awal, sehingga mengungkap struktur alami data (Putra & Fadhillah, 2025). Pendekatan aglomeratif memungkinkan visualisasi dendrogram yang memudahkan analisis hubungan antar data secara bertahap. Metode ini membantu mengidentifikasi provinsi dengan pola akses TIK serupa untuk menentukan prioritas intervensi kebijakan (Mulyana et al., 2024).

Tujuan penelitian ini adalah memetakan kesenjangan digital antar provinsi di Indonesia menggunakan *hierarchical clustering*. Penelitian ini

diharapkan memberikan kontribusi bagi perumusan kebijakan pembangunan digital yang lebih terarah dan berbasis data. Informasi yang dihasilkan dapat menjadi dasar pemerintah dalam memprioritaskan pengembangan infrastruktur digital di daerah tertinggal (Widiasanti et al., 2025).

KAJIAN TEORI

Perkembangan teknologi informasi dan komunikasi di Indonesia mengalami peningkatan yang sangat pesat dalam beberapa tahun terakhir. Transformasi digital mendorong masyarakat untuk semakin bergantung pada perangkat teknologi seperti telepon seluler, internet, dan komputer/laptop dalam menjalankan aktivitas sehari-hari, baik pada bidang pendidikan, ekonomi, komunikasi, maupun pemerintahan digital. Menurut Badan Pusat Statistik, penggunaan internet dan perangkat digital terus meningkat dari tahun ke tahun seiring berkembangnya masyarakat informasi dan kebutuhan akses data secara cepat (Badan Pusat Statistik, 2021).

Penggunaan *smartphone* merupakan salah satu bentuk pemanfaatan teknologi komunikasi yang berkembang sangat pesat di Indonesia. *Smartphone* tidak hanya digunakan sebagai alat komunikasi, tetapi juga sebagai sarana akses informasi, media pembelajaran, transaksi digital, dan aktivitas sosial masyarakat. Tingginya persentase penduduk yang menggunakan *smartphone* menunjukkan semakin meningkatnya adopsi teknologi digital di masyarakat. Penggunaan *smartphone* dipengaruhi oleh beberapa faktor, seperti tingkat pendidikan, kondisi ekonomi, infrastruktur jaringan, dan kebutuhan komunikasi masyarakat. Dalam menjelaskan proses penerimaan teknologi tersebut, teori yang relevan adalah *Diffusion of Innovation Theory* yang dikemukakan oleh Everett Rogers. Teori ini menjelaskan bahwa suatu inovasi atau teknologi akan diadopsi secara bertahap oleh masyarakat melalui proses penyebaran informasi dan penerimaan inovasi berdasarkan manfaat yang dirasakan.

Selain penggunaan *smartphone*, akses internet juga menjadi indikator penting dalam pembangunan digital. Internet merupakan jaringan komunikasi global yang memungkinkan masyarakat memperoleh informasi, melakukan komunikasi, pembelajaran daring, serta aktivitas ekonomi digital.

Persentase penduduk yang mengakses internet dalam tiga bulan terakhir menunjukkan tingkat keterhubungan masyarakat terhadap teknologi informasi. Semakin tinggi tingkat akses internet, maka semakin besar pula peluang masyarakat untuk memperoleh informasi dan meningkatkan kualitas hidup. Namun, akses internet di Indonesia masih menunjukkan adanya kesenjangan antarwilayah, terutama antara daerah perkotaan dan pedesaan. Perbedaan tersebut dipengaruhi oleh faktor ekonomi, infrastruktur jaringan, pendidikan, dan kondisi geografis. Kondisi ini dijelaskan dalam *Digital Divide Theory* yang menyatakan bahwa tidak semua masyarakat memiliki kesempatan dan kemampuan yang sama dalam mengakses teknologi digital.

Dalam penelitian ini, metode yang digunakan adalah *Hierarchical Clustering*. *Hierarchical Clustering* merupakan metode analisis cluster yang digunakan untuk mengelompokkan objek berdasarkan tingkat kemiripan karakteristik data. Metode ini membentuk struktur hierarki dalam bentuk dendrogram yang menunjukkan kedekatan antarobjek atau wilayah. Penggunaan *Hierarchical Clustering* pada data TIK dapat membantu mengelompokkan provinsi atau daerah berdasarkan karakteristik penggunaan smartphone, akses internet, dan kepemilikan komputer/laptop sehingga dapat diketahui wilayah dengan tingkat perkembangan digital yang relatif sama (Alif et al., 2024).

Namun demikian, peningkatan penggunaan internet tidak selalu diikuti pemerataan akses teknologi di seluruh wilayah Indonesia. Konsep *digital divide* atau kesenjangan digital menjelaskan adanya perbedaan akses teknologi antar kelompok masyarakat, baik berdasarkan wilayah, tingkat pendidikan, maupun kondisi ekonomi. Penelitian di Indonesia menunjukkan bahwa rumah tangga di wilayah perkotaan memiliki peluang lebih besar dalam memperoleh akses internet dibandingkan rumah tangga di wilayah pedesaan (Wahyunengseh et al., 2020).

Dalam penelitian ini digunakan metode hierarki atau analisis hierarki untuk mengelompokkan wilayah berdasarkan kemiripan karakteristik penggunaan teknologi digital. Penelitian sebelumnya menunjukkan bahwa metode *Hierarchical Clustering* efektif digunakan dalam pengelompokan data sosial dan teknologi di Indonesia. Penelitian mengenai kondisi keterampilan, akses, dan fasilitas TIK di

Indonesia menunjukkan bahwa metode *Agglomerative Ward* mampu menghasilkan pengelompokan wilayah yang optimal dalam mengidentifikasi ketimpangan digital antarwilayah (Rahmawatin et al., 2024). Selain itu, penelitian lain juga menunjukkan bahwa *Hierarchical Clustering* mampu mengelompokkan data berdasarkan karakteristik yang serupa dan memberikan visualisasi hubungan antar cluster melalui dendrogram (Alif et al., 2024).

METODE

1. Data dan Variabel Penelitian

Penelitian ini menggunakan data sekunder yang mencakup seluruh provinsi di Indonesia, yaitu sebanyak 38 provinsi pada tahun 2024. Data tersebut merepresentasikan tingkat pemanfaatan teknologi informasi dan komunikasi oleh masyarakat di masing-masing wilayah. Setiap provinsi dijadikan sebagai unit analisis, sehingga total observasi dalam penelitian ini berjumlah 38.

Variabel yang digunakan dalam penelitian ini terdiri dari tiga indikator utama yang menggambarkan akses dan penggunaan teknologi oleh masyarakat. Variabel pertama adalah persentase pengguna handphone (PctHP), yang menunjukkan proporsi penduduk di suatu provinsi yang menggunakan telepon seluler. Variabel kedua adalah persentase pengguna komputer (PctKomputer), yang menggambarkan tingkat penggunaan perangkat komputer atau laptop oleh masyarakat. Variabel ketiga adalah persentase pengguna internet (PctInternet), yang menunjukkan seberapa besar proporsi penduduk yang telah mengakses internet.

Ketiga variabel tersebut memiliki satuan yang sama dalam bentuk persentase, namun tetap dilakukan proses standardisasi menggunakan metode Z-score untuk menghindari adanya dominasi variabel tertentu dalam proses pengelompokan. Dengan standardisasi ini, setiap variabel memiliki rata-rata nol dan standar deviasi satu, sehingga analisis *clustering* menjadi lebih objektif dan seimbang.

2. Metode yang Digunakan

Metode yang digunakan dalam penelitian ini adalah analisis cluster hierarki (*hierarchical clustering*) dengan pendekatan *agglomerative*. Metode ini bekerja dengan cara mengelompokkan objek secara bertahap, dimulai dari setiap objek sebagai satu

cluster, kemudian digabungkan secara berulang hingga membentuk satu kelompok besar.

Dalam penelitian ini digunakan beberapa metode penggabungan (*linkage*), yaitu *complete linkage*, *average linkage*, dan *Ward's method*. Ketiga metode tersebut dibandingkan untuk menentukan metode terbaik dalam membentuk struktur *cluster* yang paling optimal. Kriteria pemilihan metode terbaik didasarkan pada nilai *Agglomerative Coefficient* (AC), di mana semakin mendekati nilai 1 menunjukkan struktur *cluster* yang semakin kuat.

Selain itu, dalam proses *clustering* digunakan ukuran jarak Euclidean untuk mengukur kedekatan antar objek. Visualisasi hasil *clustering* dilakukan menggunakan dendrogram, yang memberikan gambaran struktur pengelompokan antar provinsi.

3. Langkah-langkah Penelitian

Langkah-langkah penelitian adalah sebagai berikut:

- Input data*.
- Standarisasi data.
- Hitung jarak *Euclidean*.
- Hitung korelasi *cophenetic* (r_{coph}).
- Pilih metode dengan r_{coph} tertinggi.
- Penentuan k optimal.
- Hitung profil *cluster*
- Visualisasi dendrogram.

HASIL DAN PEMBAHASAN

1. Statistik Deskriptif

Sebelum analisis *clustering* dilakukan, data tiga indikator literasi digital pada 38 provinsi di Indonesia distandarisasi menggunakan *z-score*. Rata-rata nasional masing-masing indikator adalah 66,75% untuk kepemilikan telepon seluler, 18,89% untuk penggunaan komputer, dan 86,27% untuk akses internet tiga bulan terakhir. Hasil statistik deskriptif selengkapnya disajikan pada Tabel 1.

Tabel 1. Statistik Deskriptif Variabel Indikator Digital

Variabel	Mean	Std. Dev
Telepon Seluler (%)	66.75	12.26
Komputer (%)	18.89	6.67
Akses Internet 3 Bulan (%)	86.27	15.99

2. Pemilihan Metode *Linkage*

Lima metode *linkage* dievaluasi menggunakan korelasi *cophenetic* (r_{coph}) untuk menentukan metode yang paling mampu merepresentasikan struktur

jarak asli antar observasi. Hasil selengkapnya disajikan pada Tabel 2.

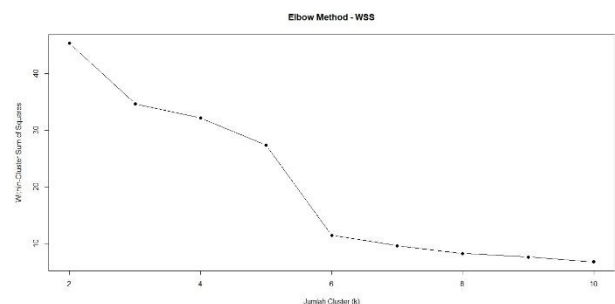
Tabel 2. Korelasi *Cophenetic*

Metode <i>Linkage</i>	r_{coph}
<i>Single</i>	0.9286
<i>Complete</i>	0.9032
<i>Average</i>	0.9243
<i>Ward.D2</i>	0.8436
<i>Centroid</i>	0.9341

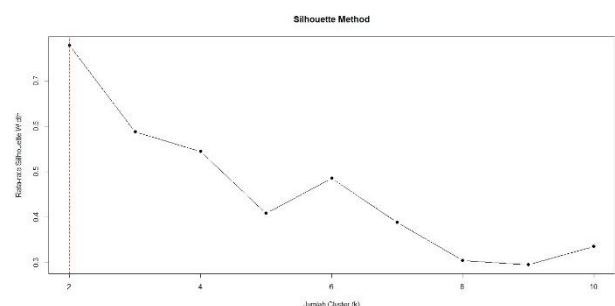
Metode *centroid* menghasilkan nilai r_{coph} tertinggi sebesar 0,9341, yang menunjukkan kesesuaian paling baik antara matriks jarak *Euclidean* dengan matriks *cophenetic* hasil dendrogram. Oleh karena itu, metode *centroid* dipilih sebagai metode *linkage* pada analisis ini.

3. Penentuan Jumlah *Cluster* Optimal

Penentuan jumlah *cluster* optimal dilakukan menggunakan tiga pendekatan, yaitu *Elbow Method* berdasarkan *Within-Cluster Sum of Squares* (WSS), *Silhouette Method*, dan *NbClust* yang mengevaluasi 26 indeks validitas *cluster* secara bersamaan. Hasil dari ketiga metode disajikan pada Tabel 3 serta Gambar 1 dan Gambar 2.



Gambar 1. Elbow-Method - Within-Cluster Sum of Squares



Gambar 2. Silhouette Method

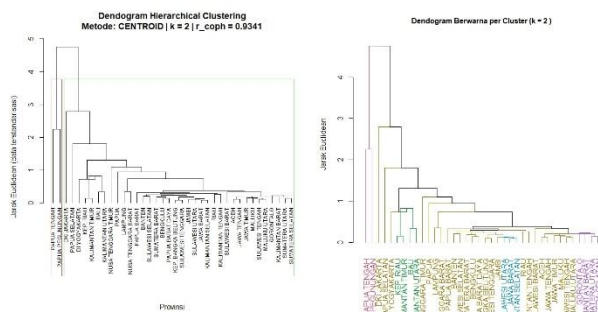
Tabel 3. Perbandingan Rekomendasi Jumlah *Cluster* Optimal

Metode	k Optimal
<i>Elbow</i> (WSS)	5
<i>Silhouette</i>	2
<i>NbClust</i>	2

Elbow Method mendeteksi titik siku pada $k = 5$ berdasarkan selisih kedua nilai WSS, sedangkan *Silhouette Method* menghasilkan rata-rata silhouette width tertinggi pada $k = 2$ sebesar 0,77. *NbClust* melalui aturan mayoritas dari 26 indeks juga merekomendasikan $k = 2$, dengan 8 dari 26 indeks menunjuk pada nilai tersebut. Berdasarkan kesepakatan dua dari tiga metode serta dikuatkan oleh nilai *silhouette* yang tinggi, $k = 2$ ditetapkan sebagai jumlah cluster final.

4. Hasil Clustering

Gambar 3 menyajikan dendrogram hasil hierarchical clustering dengan metode *centroid* pada $k = 2$. Kotak berwarna menandai dua kelompok yang terbentuk. Terlihat bahwa Papua Tengah dan Papua Pegunungan memisahkan diri secara jelas dari 36 provinsi lainnya, dengan jarak *Euclidean* yang sangat besar, yaitu lebih dari 4 satuan terstandarisasi, dibandingkan jarak antar provinsi dalam kelompok yang lebih besar.



Gambar 3. Dendrogram Hierarchical Clustering

Tabel 4 menyajikan anggota setiap cluster beserta jumlah provinsinya.

Tabel 4. Komposisi Anggota Cluster

Cluster	Jumlah Provinsi	Anggota
1	36	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat,

Maluku, Maluku Utara, Papua Barat, Papua Barat Daya, Papua, Papua Selatan, Papua Tengah, Papua Pegunungan

Rata-rata nilai asli ketiga variabel per cluster disajikan pada Tabel 5.

Tabel 5. Profil Rata-rata Variabel per Cluster

Cluster	Telepon Seluler (%)	Komputer (%)	Akses Internet 3 Bulan (%)
1	69.081	19.656	89.728
2	24.73	5.095	24.115

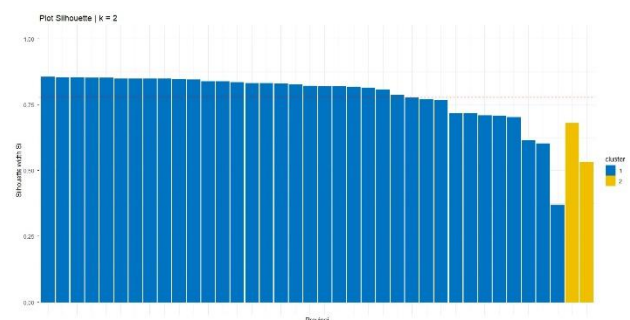
Cluster 1 memiliki rata-rata kepemilikan telepon seluler sebesar 69,08%, penggunaan komputer 19,66%, dan akses internet 89,73%. Cluster 2 memperlihatkan nilai yang jauh lebih rendah pada semua indikator, yakni 24,73% untuk telepon seluler, 5,10% untuk komputer, dan 24,12% untuk akses internet. Selisih pada indikator akses internet bahkan mencapai lebih dari 65 poin persentase, yang mengindikasikan kesenjangan digital yang sangat dalam antara kedua kelompok.

5. Evaluasi Kualitas Cluster

Gambar 4 menampilkan plot silhouette per observasi untuk $k = 2$. Hampir seluruh provinsi dalam Cluster 1 memiliki nilai silhouette di atas 0,75, dengan rata-rata 0,79. Cluster 2 memiliki rata-rata silhouette 0,61. Keduanya berada di atas ambang 0,50 yang menandakan struktur cluster yang cukup kuat hingga kuat. Ringkasan nilai silhouette per cluster disajikan pada Tabel 6.

Tabel 6. Rata-rata Silhouette Width per Cluster

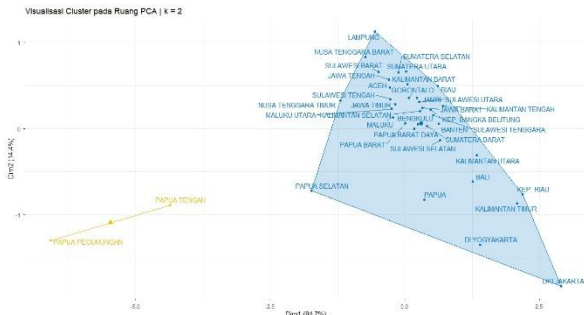
Cluster	Jumlah Provinsi	Rata-rata Silhouette Width
1	36	0.79
2	2	0.61



Gambar 4. Plot Silhouette

Gambar 5 menunjukkan proyeksi hasil clustering pada dua komponen utama pertama. Dimensi 1 menjelaskan 84,7% variansi dan Dimensi 2 menjelaskan 14,4%, sehingga total variansi yang tertangkap mencapai 99,1%. Cluster 2 terlihat terpisah

jauh di sisi kiri bidang koordinat, mengindikasikan karakteristik digital yang sangat berbeda dengan provinsi lainnya. *Cluster 1* membentuk *convex hull* yang luas di sisi kanan, mencerminkan keberagaman internal namun tetap dalam satu kelompok yang koheren.



Gambar 5. Visualisasi *Cluster* pada Ruang PCA

6. Pembahasan

Hasil analisis menunjukkan adanya kesenjangan digital yang sangat nyata antara dua kelompok provinsi di Indonesia. *Cluster 1* yang beranggotakan 36 provinsi mencakup seluruh wilayah di luar Papua pedalaman, dengan tingkat penetrasi digital yang relatif baik meskipun terdapat variasi di dalamnya. Provinsi seperti DKI Jakarta, Kep. Riau, dan Kalimantan Timur memiliki *z-score* yang jauh di atas rata-rata nasional, sementara provinsi seperti Nusa Tenggara Timur dan Lampung berada di bawah rata-rata namun masih jauh lebih tinggi dibandingkan *Cluster 2*.

Cluster 2, yang hanya terdiri dari Papua Tengah dan Papua Pegunungan, merupakan *outlier* ekstrem dalam lanskap digital Indonesia. Nilai *z-score* Papua Pegunungan pada variabel telepon seluler mencapai -4,16 dan akses internet -4,64, yang berarti kondisinya berada lebih dari empat standar deviasi di bawah rata-rata nasional. Kondisi ini mencerminkan tantangan infrastruktur yang sangat berat, termasuk keterbatasan jaringan telekomunikasi, topografi pegunungan yang sulit dijangkau, serta ketersediaan listrik yang terbatas di wilayah tersebut.

Tingginya nilai korelasi *cophenetic* pada metode centroid ($r_{coph} = 0,9341$) menunjukkan bahwa dendrogram yang dihasilkan merepresentasikan struktur jarak antar provinsi dengan sangat baik. Rata-rata *silhouette* keseluruhan yang tinggi pada *Cluster 1* sebesar 0,79 mengkonfirmasi bahwa pemisahan dua cluster ini bukan artefak algoritmik, melainkan mencerminkan perbedaan kondisi nyata

yang substansial. Temuan ini memperkuat urgensi kebijakan afirmatif di bidang infrastruktur digital, khususnya di wilayah Papua pedalaman, sebagai prasyarat bagi pemerataan akses terhadap layanan pendidikan, kesehatan, dan ekonomi digital di seluruh wilayah Indonesia.

PENUTUP

SIMPULAN

Berdasarkan hasil analisis yang telah dilakukan, dapat disimpulkan bahwa metode hierarchical clustering dengan pendekatan centroid merupakan metode terbaik dalam mengelompokkan provinsi di Indonesia berdasarkan indikator kesenjangan digital. Hal ini ditunjukkan oleh nilai korelasi *cophenetic* tertinggi dibandingkan empat metode linkage lainnya, yaitu sebesar 0,9341, yang mengindikasikan bahwa dendrogram yang terbentuk merepresentasikan struktur jarak antar provinsi dengan sangat baik.

Berdasarkan tiga metode penentuan jumlah cluster optimal, yaitu *Elbow Method*, *Silhouette Method*, dan *NbClust*, diperoleh kesepakatan pada $k = 2$ sebagai jumlah cluster final. Hasil clustering menunjukkan bahwa terdapat kesenjangan digital yang sangat nyata antar provinsi di Indonesia. *Cluster 1* beranggotakan 36 provinsi dengan rata-rata akses internet sebesar 89,73%, kepemilikan telepon seluler 69,08%, dan penggunaan komputer 19,66%. *Cluster 2* hanya terdiri dari Papua Tengah dan Papua Pegunungan, dengan nilai yang jauh lebih rendah pada seluruh indikator, yaitu akses internet 24,12%, telepon seluler 24,73%, dan komputer 5,10%. Selisih akses internet antar kedua cluster mencapai lebih dari 65 poin persentase, yang mencerminkan kesenjangan infrastruktur digital yang sangat dalam di wilayah Papua pedalaman dibandingkan wilayah lainnya di Indonesia.

Kualitas cluster yang terbentuk dikonfirmasi melalui rata-rata *silhouette width* sebesar 0,79 untuk *Cluster 1* dan 0,61 untuk *Cluster 2*, yang keduanya berada di atas ambang 0,50 dan menandakan struktur cluster yang kuat.

SARAN

Berdasarkan hasil penelitian yang diperoleh, terdapat beberapa saran yang dapat diberikan. Dari

sisi kebijakan, pemerintah diharapkan memberikan perhatian lebih terhadap wilayah yang memiliki tingkat penetrasi digital sangat rendah, khususnya di Papua Tengah dan Papua Pegunungan. Upaya yang dapat dilakukan meliputi peningkatan infrastruktur jaringan telekomunikasi, penyediaan perangkat teknologi yang terjangkau, perluasan jaringan listrik, serta peningkatan literasi digital masyarakat di wilayah tersebut.

Dari sisi metodologi, penelitian selanjutnya disarankan untuk menambahkan variabel yang lebih beragam dan relevan, seperti tingkat pendidikan, tingkat pendapatan masyarakat, serta ketersediaan infrastruktur digital, agar hasil clustering dapat memberikan gambaran kesenjangan digital yang lebih menyeluruh. Selain itu, penggunaan metode *clustering* lain seperti K-Means atau DBSCAN juga dapat dipertimbangkan untuk membandingkan hasil dan meningkatkan validitas penelitian.

kondisi keterampilan, akses dan fasilitas teknologi informasi dan komunikasi di Indonesia. *Komputika: Jurnal Sistem Komputer*, 13(1), 83-92.

Wahyunengseh, R. D., Hastjarjo, S., Mulyaningsih, T., & Suharto, D. G. (2020). Digital governance and digital divide: A matrix of the poor's vulnerabilities. *Policy & Governance Review*, 4(2), 152-166.

Widiasanti, I., Rahmadani, S., Az-Zahra Nur, D., Putri Lestari, N., & Syaidah, S. (2025). Kesetaraan akses internet dan tantangan literasi digital di Indonesia. *Jurnal Pendidikan Tambusai*, 9(2), 19631-19637

DAFTAR PUSTAKA

- Alif, M. N., Huda, M., & Rahmawati, A. (2024). Penerapan metode hierarchical clustering untuk pengelompokan data teknologi informasi dan komunikasi di Indonesia. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(2), 245-254.
- Badan Pusat Statistik. (2021). *Statistik telekomunikasi Indonesia 2021*. BPS-Statistics Indonesia.
- Jayanthi, R., & Dinaseviani, A. (2022). Kesenjangan digital dan solusi yang diterapkan di Indonesia selama pandemi COVID-19. *Jurnal IPTEKKOM: Jurnal Ilmu Pengetahuan & Teknologi Informasi*, 24(2), 187-200. <https://doi.org/10.17933/iptekkom.24.2.2022.187-200>.
- Mulyana, A. A. R., Juwita, A. R., Siregar, A. M., & Fauzi, A. (2024). Penerapan algoritma K-means clustering dan hierarchical clustering dalam mengelompokkan data pengangguran di Karawang. *Jurnal Algoritma*, 21(2), 209-220. <https://doi.org/10.33364/algoritma/v.21-2.2155>.
- Putra, I. P., & Fadhillah, A. (2025). Perbandingan metode K-Means dan hierarchical clustering dalam pengelompokan data penduduk miskin di Kabupaten Cianjur. *LANCAH: Jurnal Inovasi dan Tren*, 3(1). <https://doi.org/10.35870/ljit.v3i1.4028>.
- Rahmawatin, N., Permatasari, N. P., Wijayanto, A. W., & Marsisno, W. (2024). Analisis cluster