



Pemetaan Kesenjangan Digital Provinsi Di Indonesia Menggunakan *Algoritma K-Means* Berdasarkan Indikator Teknologi Informasi Dan Komunikasi Tahun 2024

Cut Asiana Gemawaty¹, Yossi Indrawati Syuhardi², Baihaqi³

¹ Sistem Informasu, Universitas Gunadarma

^{2,3} Teknik Informatika, Universitas Indraprasta PGRI

¹cut_asiana@staff.gunadarma.ac.id, ²yossiindrawatisyuhardi@gmail.com, ³bay.haqi@unindra.ac.id

Abstrak

Perkembangan teknologi informasi dan komunikasi (TIK) memiliki peran penting dalam mendukung transformasi digital di berbagai sektor. Namun, tingkat akses dan pemanfaatan TIK antarprovinsi di Indonesia masih menunjukkan perbedaan yang cukup signifikan. Penelitian ini bertujuan untuk mengelompokkan provinsi di Indonesia berdasarkan indikator teknologi informasi dan komunikasi menggunakan algoritma *K-Means Clustering* sehingga diperoleh pemetaan tingkat digitalisasi provinsi yang dapat digunakan sebagai dasar perumusan kebijakan pemerataan pembangunan digital. Penelitian menggunakan data sekunder yang bersumber dari *Statistik Telekomunikasi Indonesia Tahun 2024* yang diterbitkan oleh Badan Pusat Statistik (BPS). Variabel yang digunakan meliputi persentase penduduk mengakses internet, persentase penduduk memiliki atau menguasai telepon seluler, persentase rumah tangga memiliki atau menguasai komputer, persentase rumah tangga mengakses internet, dan persentase rumah tangga memiliki atau menguasai telepon seluler. Tahapan penelitian meliputi *preprocessing* data, standarisasi menggunakan metode *Z-score*, penentuan jumlah cluster optimal menggunakan *Elbow Method*, pengelompokan menggunakan algoritma *K-Means*, serta evaluasi menggunakan *Silhouette Score*. Hasil penelitian menunjukkan bahwa jumlah cluster optimal adalah tiga cluster. Cluster digital tinggi terdiri atas enam provinsi, cluster digital menengah terdiri atas 31 provinsi, dan cluster digital rendah terdiri atas dua provinsi, yaitu Papua Tengah dan Papua Pegunungan. Hasil penelitian menunjukkan bahwa kesenjangan digital antarprovinsi di Indonesia masih cukup signifikan. Temuan ini diharapkan dapat menjadi masukan bagi pemerintah dalam menyusun kebijakan pemerataan pembangunan digital yang lebih efektif dan berkelanjutan.

Kata Kunci: *K-Means Clustering*, Teknologi Informasi Dan Komunikasi, Kesenjangan Digital, Digitalisasi Provinsi, *Data Mining*.

1. Pendahuluan

Perkembangan teknologi informasi dan komunikasi (TIK) telah menjadi salah satu faktor utama yang mendorong transformasi ekonomi, sosial, dan pendidikan di era digital. TIK berperan penting dalam meningkatkan produktivitas, efisiensi, dan inovasi di berbagai sektor melalui pemanfaatan teknologi digital yang semakin luas. Dalam bidang ekonomi, perkembangan TIK telah mendorong pertumbuhan ekonomi digital melalui berbagai aktivitas seperti *electronic commerce (e-commerce)*, *digital marketing*, dan *cloud computing services* yang mampu meningkatkan konektivitas serta memperluas akses pasar secara global (Xia et al., 2024). Selain itu, TIK juga merevolusi pola komunikasi masyarakat dengan memungkinkan pertukaran informasi secara cepat, kolaboratif, dan tanpa batas geografis, sehingga mendukung partisipasi sosial dan ekonomi dalam masyarakat berbasis informasi (Warschauer & Matuchniak, 2010). Pada sektor pendidikan, pemanfaatan TIK berkontribusi dalam memperluas akses terhadap sumber belajar, meningkatkan kualitas pembelajaran, serta mendukung pengembangan kompetensi digital yang dibutuhkan pada abad ke-21 (Roztock et al., 2019). Berbagai penelitian menunjukkan bahwa TIK telah menjadi pendorong penting pembangunan berkelanjutan dan transformasi digital yang memengaruhi hampir seluruh aspek kehidupan masyarakat modern (Oh et al., 2018).

Meskipun transformasi digital di Indonesia terus berkembang, tingkat pemanfaatan teknologi digital antarwilayah masih menunjukkan kesenjangan yang cukup signifikan. Berbagai penelitian menunjukkan bahwa perkembangan digital lebih terkonsentrasi di wilayah maju, khususnya Pulau Jawa, dibandingkan dengan wilayah Indonesia bagian timur yang masih menghadapi keterbatasan infrastruktur dan akses teknologi digital (Jaya et al., 2024; Kartiasih et al., 2023a). Selain itu, perbedaan keterampilan digital, kondisi sosial ekonomi, dan akses terhadap perangkat teknologi turut memengaruhi tingkat adopsi teknologi di setiap daerah (Jaya et al., 2026). Kondisi tersebut menunjukkan bahwa pemerataan pembangunan digital masih menjadi tantangan yang perlu mendapat perhatian guna mewujudkan transformasi digital yang inklusif dan berkelanjutan di Indonesia.

Kesenjangan digital antarprovinsi dipengaruhi oleh berbagai faktor, seperti tingkat pembangunan ekonomi, kondisi geografis, ketersediaan infrastruktur telekomunikasi, serta karakteristik sosial ekonomi masyarakat. Di

Pemetaan Kesenjangan Digital Provinsi Di Indonesia Menggunakan *Algoritma K-Means* Berdasarkan Indikator Teknologi Informasi Dan Komunikasi Tahun 2024

Indonesia, perkembangan digital cenderung lebih tinggi pada wilayah metropolitan dan provinsi di bagian barat dibandingkan dengan wilayah perdesaan dan kawasan timur Indonesia yang masih menghadapi keterbatasan akses teknologi digital (Kartiasih et al., 2023). Selain itu, faktor pendidikan, pendapatan, dan kesempatan kerja turut memengaruhi tingkat adopsi dan pemanfaatan teknologi informasi dan komunikasi. Kondisi tersebut menunjukkan bahwa kesenjangan digital tidak hanya berkaitan dengan akses terhadap teknologi, tetapi juga mencerminkan perbedaan kapasitas sosial dan ekonomi antarwilayah yang perlu mendapat perhatian dalam upaya mewujudkan pemerataan pembangunan digital.

Data Statistik Telekomunikasi Indonesia Tahun 2024 yang diterbitkan oleh Badan Pusat Statistik (BPS) menunjukkan adanya variasi yang cukup besar pada indikator teknologi informasi dan komunikasi antarprovinsi. Perbedaan tersebut terlihat pada tingkat akses internet penduduk, kepemilikan telepon seluler, kepemilikan komputer rumah tangga, serta akses internet rumah tangga. Beberapa provinsi seperti DKI Jakarta, Kepulauan Riau, dan Bali memiliki tingkat pemanfaatan TIK yang relatif tinggi, sedangkan beberapa provinsi di kawasan timur Indonesia masih menunjukkan tingkat akses dan pemanfaatan TIK yang lebih rendah. Kondisi ini mengindikasikan bahwa pembangunan digital di Indonesia belum berlangsung secara merata sehingga diperlukan upaya pemetaan untuk mengidentifikasi karakteristik tingkat digitalisasi pada masing-masing provinsi.

Identifikasi wilayah berdasarkan tingkat perkembangan TIK menjadi penting untuk mendukung perumusan kebijakan yang bertujuan mengurangi kesenjangan digital dan meningkatkan pemerataan pembangunan digital. Salah satu pendekatan yang dapat digunakan adalah teknik *clustering* yang mampu mengelompokkan wilayah berdasarkan tingkat kemiripan karakteristik indikator TIK. Berbagai penelitian menunjukkan bahwa metode *clustering* efektif digunakan untuk mengidentifikasi tingkat perkembangan digital suatu wilayah dan mengungkap pola kesenjangan digital yang terjadi (Ayanso et al., 2010, 2014). Oleh karena itu, penelitian ini menerapkan algoritma *K-Means Clustering* untuk mengelompokkan provinsi di Indonesia berdasarkan indikator teknologi informasi dan komunikasi tahun 2024. Hasil pengelompokan diharapkan dapat memberikan gambaran mengenai tingkat digitalisasi masing-masing provinsi serta menjadi masukan bagi pemerintah dalam menyusun kebijakan pemerataan pembangunan digital yang lebih tepat sasaran.

Algoritma *K-Means Clustering* telah banyak digunakan dalam berbagai penelitian untuk mengelompokkan wilayah berdasarkan indikator sosial, ekonomi, dan pembangunan. Simbolon et al. (2025) menerapkan algoritma *K-Means* untuk mengelompokkan wilayah berdasarkan tingkat kemiskinan di Provinsi Papua dan berhasil mengidentifikasi kelompok daerah dengan karakteristik kemiskinan yang berbeda. Penelitian serupa juga dilakukan pada wilayah Kepulauan Maluku dan Papua yang menghasilkan tiga kelompok kabupaten berdasarkan tingkat kemiskinan sehingga dapat mendukung penyusunan kebijakan pembangunan yang lebih terarah (Matdoan et al., 2024). Selain itu, metode *K-Means* telah digunakan untuk mengelompokkan daerah berdasarkan tingkat kemiskinan di Provinsi Banten dan Sumatera Utara serta berdasarkan Indeks Pembangunan Manusia (IPM) di Jawa Tengah (Andriyani et al., 2023; Fadillah et al., 2025; Handayanna, 2023). Berbagai penelitian tersebut menunjukkan bahwa algoritma *K-Means* merupakan metode yang efektif dalam mengidentifikasi karakteristik wilayah dan mendukung pengambilan keputusan berbasis data dalam perencanaan pembangunan daerah.

Meskipun penelitian mengenai pengelompokan wilayah menggunakan algoritma *K-Means* telah banyak dilakukan, sebagian besar penelitian masih berfokus pada indikator kemiskinan, pembangunan manusia, dan pembangunan wilayah. Penelitian yang secara khusus mengelompokkan provinsi di Indonesia berdasarkan indikator teknologi informasi dan komunikasi menggunakan data terbaru dari Statistik Telekomunikasi Indonesia Tahun 2024 masih relatif terbatas. Oleh karena itu, penelitian ini mengisi kesenjangan tersebut dengan menerapkan algoritma *K-Means Clustering* untuk memetakan tingkat digitalisasi provinsi di Indonesia berdasarkan lima indikator TIK, yaitu akses internet penduduk, kepemilikan telepon seluler penduduk, kepemilikan komputer rumah tangga, akses internet rumah tangga, dan kepemilikan telepon seluler rumah tangga.

Penelitian ini bertujuan untuk mengelompokkan provinsi di Indonesia berdasarkan indikator teknologi informasi dan komunikasi menggunakan algoritma *K-Means Clustering*. Pengelompokan dilakukan untuk mengidentifikasi karakteristik tingkat digitalisasi setiap provinsi sehingga diperoleh pemetaan kesenjangan digital antarwilayah di Indonesia. Hasil penelitian diharapkan dapat menjadi masukan bagi pemerintah dan pemangku kepentingan dalam merumuskan kebijakan yang mendukung pemerataan pembangunan digital secara lebih efektif dan berkelanjutan.

2. Metode Penelitian

2.1 Data Penelitian

Penelitian ini menggunakan data sekunder yang diperoleh dari publikasi Statistik Telekomunikasi Indonesia Tahun 2024 yang diterbitkan oleh Badan Pusat Statistik (BPS). Data yang digunakan merupakan data tingkat provinsi tahun 2024 dengan jumlah objek penelitian sebanyak 38 provinsi di Indonesia. Data tersebut dipilih karena mampu merepresentasikan kondisi perkembangan teknologi informasi dan komunikasi (TIK) di setiap provinsi berdasarkan indikator yang relevan dengan tingkat digitalisasi masyarakat.

2.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri atas lima indikator teknologi informasi dan komunikasi yang disajikan pada Tabel 1.

Tabel 1. Variabel Penelitian

Kode	Variabel
X1	Persentase Penduduk Mengakses Internet
X2	Persentase Penduduk Memiliki/Menguasai Telepon Seluler
X3	Persentase Rumah Tangga Memiliki/Menguasai Komputer
X4	Persentase Rumah Tangga Mengakses Internet
X5	Persentase Rumah Tangga Memiliki/Menguasai Telepon Seluler

Sumber: Statistik Telekomunikasi Indonesia 2024 (BPS), diolah penulis.

2.3 Tahapan Penelitian

Tahapan penelitian dimulai dengan pengumpulan data dari publikasi BPS, dilanjutkan dengan *preprocessing* data untuk memastikan kelengkapan dan konsistensi data. Selanjutnya dilakukan standarisasi data menggunakan metode *Z-Score* untuk menyamakan skala antarvariabel. Tahap berikutnya adalah penentuan jumlah cluster optimal menggunakan *Elbow Method*, kemudian dilakukan proses pengelompokan menggunakan algoritma *K-Means Clustering*. Hasil pengelompokan selanjutnya dievaluasi menggunakan *Silhouette Score* dan diinterpretasikan berdasarkan karakteristik masing-masing cluster.

2.4 Standardisasi Data

Karena setiap variabel memiliki rentang nilai yang berbeda, maka dilakukan proses standarisasi menggunakan metode *Z-Score*. Standarisasi bertujuan untuk menghilangkan pengaruh perbedaan skala antarvariabel sehingga setiap variabel memiliki kontribusi yang seimbang dalam proses pengelompokan.

Rumus standarisasi *Z-Score* ditunjukkan pada Persamaan (1).

$$Z = \frac{x - \mu}{\sigma}$$

Keterangan:

Z = nilai hasil standarisasi (*Z-score*)

x = nilai data asli

μ = nilai rata-rata (*mean*) variabel

σ = standar deviasi (*standard deviation*) variabel

2.5 Algoritma K-Means Clustering

Algoritma *K-Means Clustering* merupakan salah satu metode *unsupervised learning* yang digunakan untuk mengelompokkan objek berdasarkan tingkat kemiripan karakteristik. Proses pengelompokan dilakukan melalui beberapa tahapan, yaitu:

1. Menentukan jumlah cluster ((K)).
2. Menentukan centroid awal secara acak.
3. Menghitung jarak setiap data terhadap centroid menggunakan jarak Euclidean.
4. Menempatkan data ke dalam cluster dengan jarak terdekat.
5. Menghitung kembali centroid berdasarkan rata-rata anggota cluster.
6. Mengulangi proses hingga tidak terjadi perubahan anggota cluster atau centroid telah konvergen.

Perhitungan jarak *Euclidean* dilakukan menggunakan Persamaan (2).

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Keterangan:

d(x,y) = jarak antara data dan centroid

x_i = nilai atribut ke-i pada data

y_i = nilai atribut ke-i pada centroid

n = jumlah atribut yang digunakan dalam proses clustering

2.6 Elbow Method

Elbow Method digunakan untuk menentukan jumlah cluster optimal dengan mengamati nilai *Sum of Squared Error* (SSE) pada beberapa alternatif jumlah cluster. Jumlah cluster optimal ditentukan berdasarkan titik siku (*elbow point*) pada grafik SSE, yaitu titik ketika penurunan nilai SSE mulai melambat secara signifikan.

Rumus SSE ditunjukkan pada Persamaan (3).

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} (x - \mu_i)^2$$

Keterangan:

SSE = Sum of Squared Error

C_i = cluster ke-i

μ_i = centroid cluster ke-i

2.7 Silhouette Score

Evaluasi kualitas cluster dilakukan menggunakan *Silhouette Score*. Metode ini digunakan untuk mengukur tingkat kemiripan suatu objek terhadap cluster tempatnya berada dibandingkan dengan cluster lainnya. Nilai *Silhouette Score* berada pada rentang -1 hingga 1, di mana nilai yang semakin mendekati 1 menunjukkan kualitas cluster yang semakin baik.

Rumus *Silhouette Score* ditunjukkan pada Persamaan (4).

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Keterangan:

S(i) = nilai silhouette untuk objek ke-i

a(i) = rata-rata jarak objek ke seluruh anggota dalam cluster yang sama

b(i) = rata-rata jarak objek ke cluster terdekat lainnya

Semakin tinggi nilai *Silhouette Score*, semakin baik kualitas pengelompokan yang dihasilkan oleh algoritma *K-Means Clustering*.

3. Hasil dan Pembahasan

3.1 Statistik Deskriptif

Statistik deskriptif digunakan untuk memberikan gambaran umum mengenai karakteristik data penelitian yang terdiri atas lima indikator teknologi informasi dan komunikasi pada 38 provinsi di Indonesia. Hasil statistik deskriptif meliputi nilai rata-rata (*mean*), standar deviasi, nilai minimum, dan nilai maksimum yang disajikan pada Tabel 2.

Tabel 2. Statistik Deskriptif Variabel Penelitian

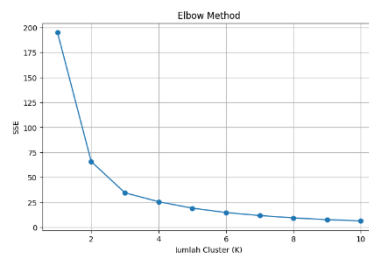
Variabel	Mean	Std. Dev	Min	Max
X1	68.16	14.65	6.76	89.26
X2	66.80	12.10	15.76	83.90
X3	18.88	6.58	2.31	39.33
X4	86.36	15.79	12.15	98.18
X5	91.33	12.91	31.66	99.15

Sumber: Hasil olah data menggunakan Python (2025).

Berdasarkan Tabel 2, variabel X5 memiliki nilai rata-rata tertinggi sebesar 91,33%, sedangkan variabel X3 memiliki nilai rata-rata terendah sebesar 18,88%. Hasil ini menunjukkan bahwa kepemilikan telepon seluler rumah tangga di Indonesia relatif tinggi, sementara kepemilikan komputer rumah tangga masih relatif rendah. Selain itu, terdapat variasi yang cukup besar antarprovinsi yang ditunjukkan oleh nilai minimum dan maksimum pada masing-masing variabel.

3.2 Penentuan Jumlah Cluster Menggunakan *Elbow Method*

Penentuan jumlah cluster optimal dilakukan menggunakan *Elbow Method*. Metode ini bertujuan untuk mengidentifikasi jumlah cluster yang paling sesuai dengan mengamati perubahan nilai *Sum of Squared Error* (SSE) pada setiap jumlah cluster yang diuji. Hasil *Elbow Method* ditunjukkan pada Gambar 2.



Gambar 2. Hasil Elbow Method

DOI: <https://doi.org/10.69693/ijmst.v4i2.10576>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Berdasarkan Gambar 2, terjadi penurunan nilai SSE yang cukup signifikan hingga K=3. Setelah titik tersebut, penurunan SSE cenderung melandai sehingga membentuk pola siku (*elbow point*). Oleh karena itu, jumlah cluster optimal yang digunakan dalam penelitian ini adalah tiga cluster.

3.3 Evaluasi Menggunakan *Silhouette Score*

Setelah jumlah cluster optimal diperoleh menggunakan *Elbow Method*, kualitas pengelompokan dievaluasi menggunakan *Silhouette Score*. Nilai *Silhouette Score* digunakan untuk mengukur tingkat kemiripan objek dalam suatu cluster dibandingkan dengan cluster lainnya. Hasil evaluasi ditampilkan pada Tabel 3.

Tabel 3. Hasil Evaluasi Menggunakan *Silhouette Score*

Jumlah Cluster (K)	<i>Silhouette Score</i>
2	0.8038
3	0.5662
4	0.3861
5	0.3272
6	0.3053

Sumber: Hasil olah data menggunakan Python (2025).

Berdasarkan Tabel 3, nilai *Silhouette Score* tertinggi diperoleh pada K=2 sebesar 0,8038. Namun demikian, hasil *Elbow Method* menunjukkan bahwa jumlah cluster optimal berada pada K=3. Selain itu, nilai *Silhouette Score* sebesar 0,5662 pada K=3 masih menunjukkan kualitas pengelompokan yang baik sehingga penelitian ini menggunakan tiga cluster dalam proses pengelompokan provinsi berdasarkan indikator TIK.

3.4 Hasil Pengelompokan Menggunakan *Algoritma K-Means*

Setelah jumlah cluster optimal ditentukan sebanyak tiga cluster, proses pengelompokan dilakukan menggunakan algoritma *K-Means Clustering*. Hasil pengelompokan menghasilkan distribusi jumlah provinsi yang berbeda pada setiap cluster. Jumlah anggota masing-masing cluster disajikan pada Tabel 4.

Tabel 4. Jumlah Anggota Cluster

Cluster	Jumlah Provinsi
0	31
1	2
2	6

Sumber: Hasil olah data menggunakan Python (2025).

Berdasarkan Tabel 4, Cluster 0 merupakan kelompok terbesar dengan jumlah 31 provinsi atau sekitar 79,49% dari total provinsi yang diamati. Sementara itu, Cluster 2 terdiri atas 6 provinsi dan Cluster 1 hanya terdiri atas 2 provinsi. Hasil tersebut menunjukkan bahwa sebagian besar provinsi di Indonesia memiliki karakteristik tingkat digitalisasi yang relatif serupa dan berada pada kelompok menengah.

3.5 Karakteristik Cluster Berdasarkan Nilai Centroid

Karakteristik masing-masing cluster dianalisis menggunakan nilai centroid yang merepresentasikan nilai rata-rata setiap indikator TIK pada setiap cluster. Nilai centroid digunakan untuk mengidentifikasi tingkat digitalisasi setiap kelompok provinsi. Hasil centroid ditunjukkan pada Tabel 5.

Tabel 5. Nilai Centroid Setiap Cluster

Cluster	X1	X2	X3	X4	X5
0	68.59	67.10	17.51	88.79	93.60
1	17.74	24.73	5.09	24.11	39.93
2	82.77	79.23	30.57	94.57	96.75

Sumber: Hasil olah data menggunakan Python (2025).

Berdasarkan Tabel 5, Cluster 2 memiliki nilai rata-rata tertinggi pada seluruh indikator TIK, yaitu akses internet penduduk, kepemilikan telepon seluler, kepemilikan komputer rumah tangga, akses internet rumah tangga, dan kepemilikan telepon seluler rumah tangga. Oleh karena itu, Cluster 2 dikategorikan sebagai kelompok dengan tingkat digitalisasi tinggi. Sebaliknya, Cluster 1 memiliki nilai rata-rata terendah pada seluruh indikator sehingga dikategorikan sebagai kelompok dengan tingkat digitalisasi rendah. Sementara itu, Cluster 0 memiliki nilai indikator yang berada di antara kedua cluster tersebut dan dikategorikan sebagai kelompok dengan tingkat digitalisasi menengah.

3.6 Pemetaan Tingkat Digitalisasi Provinsi di Indonesia

Berdasarkan hasil analisis centroid, masing-masing cluster kemudian diberi label digitalisasi tinggi, menengah, dan rendah. Distribusi provinsi pada setiap cluster disajikan pada Tabel 6.

Tabel 6. Hasil Pengelompokan Provinsi Berdasarkan Tingkat Digitalisasi

DOI: <https://doi.org/10.69693/ijmst.v4i2.10576>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Kategori	Provinsi
Digital Tinggi	Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Bali, Kalimantan Timur, Kalimantan Utara
Digital Menengah	31 provinsi lainnya
Digital Rendah	Papua Tengah, Papua Pegunungan

Sumber: Hasil olah data menggunakan Python (2025).

Hasil pengelompokan menunjukkan bahwa provinsi yang termasuk dalam kategori digital tinggi didominasi oleh wilayah dengan tingkat pembangunan dan infrastruktur TIK yang relatif maju. Sebaliknya, Papua Tengah dan Papua Pegunungan berada pada kategori digital rendah karena memiliki nilai indikator TIK yang jauh lebih rendah dibandingkan provinsi lainnya. Temuan ini menunjukkan bahwa kesenjangan digital antarprovinsi di Indonesia masih cukup nyata dan memerlukan perhatian khusus dalam penyediaan infrastruktur telekomunikasi, akses internet, serta peningkatan literasi digital masyarakat.

3.6 Pembahasan

Hasil pengelompokan menggunakan algoritma *K-Means Clustering* menunjukkan adanya tiga kategori tingkat digitalisasi provinsi di Indonesia, yaitu digital tinggi, digital menengah, dan digital rendah. *Cluster* digital tinggi terdiri atas Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Bali, Kalimantan Timur, dan Kalimantan Utara. Provinsi-provinsi tersebut memiliki nilai rata-rata tertinggi pada seluruh indikator TIK yang digunakan dalam penelitian. Tingginya tingkat digitalisasi pada wilayah tersebut dapat dikaitkan dengan ketersediaan infrastruktur telekomunikasi yang lebih baik, tingkat urbanisasi yang tinggi, serta aktivitas ekonomi yang lebih berkembang dibandingkan wilayah lainnya.

Sebagian besar provinsi di Indonesia tergolong ke dalam *cluster* digital menengah. Kondisi ini menunjukkan bahwa akses dan pemanfaatan teknologi informasi dan komunikasi telah berkembang cukup baik, namun belum mencapai tingkat optimal sebagaimana yang ditunjukkan oleh provinsi pada *cluster* digital tinggi. Hasil ini mengindikasikan bahwa pembangunan digital di Indonesia masih terkonsentrasi pada beberapa wilayah tertentu sehingga diperlukan upaya pemerataan infrastruktur dan layanan digital agar manfaat transformasi digital dapat dirasakan secara lebih luas.

Sementara itu, Papua Tengah dan Papua Pegunungan berada pada *cluster* digital rendah dengan nilai indikator TIK yang paling rendah dibandingkan *cluster* lainnya. Temuan ini sejalan dengan penelitian Kartiasih et al., (2023) yang menyatakan bahwa perkembangan digital di Indonesia masih menunjukkan kesenjangan antara wilayah barat dan timur. Keterbatasan infrastruktur telekomunikasi, kondisi geografis yang menantang, serta rendahnya akses terhadap perangkat digital menjadi faktor yang berkontribusi terhadap rendahnya tingkat digitalisasi di kedua provinsi tersebut.

Temuan penelitian ini juga memperkuat hasil penelitian Jaya et al. (2024) yang menunjukkan bahwa transformasi digital di Indonesia masih terkonsentrasi pada wilayah yang memiliki tingkat pembangunan ekonomi dan infrastruktur yang lebih baik. Oleh karena itu, kebijakan pemerataan pembangunan digital perlu difokuskan pada wilayah yang berada dalam kategori digital rendah melalui peningkatan akses internet, pembangunan infrastruktur telekomunikasi, serta penguatan literasi digital masyarakat. Dengan demikian, kesenjangan digital antarprovinsi dapat dikurangi sehingga mendukung terwujudnya transformasi digital yang inklusif dan berkelanjutan.

4. Kesimpulan

Penelitian ini berhasil mengelompokkan 38 provinsi di Indonesia berdasarkan lima indikator teknologi informasi dan komunikasi menggunakan algoritma *K-Means Clustering*. Hasil analisis menggunakan *Elbow Method* menunjukkan bahwa jumlah cluster optimal adalah tiga cluster, yang selanjutnya dikategorikan sebagai digital tinggi, digital menengah, dan digital rendah. Cluster digital tinggi terdiri atas enam provinsi, yaitu Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Bali, Kalimantan Timur, dan Kalimantan Utara, yang memiliki nilai indikator TIK tertinggi dibandingkan provinsi lainnya. Sebagian besar provinsi berada pada cluster digital menengah dengan jumlah 31 provinsi, sedangkan cluster digital rendah terdiri atas Papua Tengah dan Papua Pegunungan yang menunjukkan nilai indikator TIK paling rendah. Hasil penelitian ini mengindikasikan bahwa kesenjangan digital antarprovinsi di Indonesia masih cukup signifikan, terutama antara wilayah yang memiliki infrastruktur dan akses teknologi yang baik dengan wilayah yang masih mengalami keterbatasan akses digital. Temuan ini dapat menjadi dasar bagi pemerintah dan pemangku kepentingan dalam merumuskan kebijakan pemerataan pembangunan digital, khususnya pada wilayah dengan tingkat digitalisasi yang masih rendah. Penelitian selanjutnya dapat mempertimbangkan penggunaan data multiyear serta penambahan indikator ekonomi, pendidikan, dan infrastruktur untuk memperoleh gambaran yang lebih komprehensif mengenai perkembangan digital di Indonesia.

Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada Badan Pusat Statistik (BPS) yang telah menyediakan data melalui publikasi *Statistik Telekomunikasi Indonesia Tahun 2024* sehingga penelitian ini dapat dilaksanakan dengan baik. Penulis juga menyampaikan apresiasi kepada Universitas Indraprasta PGRI atas dukungan akademik yang diberikan selama proses penelitian dan penyusunan artikel ini.

Reference

- Andriyani, W., Nasyuha, A. H., Syahra, Y., & Triaji, B. (2023). Clustering Analysis Of Poverty Levels In North Sumatra Province Using The Application Of Data Mining With The K-Means Algorithm. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(4), 1971. <https://doi.org/10.30865/Mib.V7i4.6867>
- Ayanso, A., Cho, D. I., & Lertwachara, K. (2010). The Digital Divide: Global And Regional ICT Leaders And Followers. *Information Technology For Development*, 16(4), 304–319. <https://doi.org/10.1080/02681102.2010.504698>
- Ayanso, A., Cho, D. I., & Lertwachara, K. (2014). Information And Communications Technology Development And The Digital Divide: A Global And Regional Assessment. *Information Technology For Development*, 20(1), 60–77. <https://doi.org/10.1080/02681102.2013.797378>
- Fadillah, M. R., Ardiansyah, M. R., Junaidi, M. R., & Fauzi, F. (2025). Penerapan K-Means Clustering Dalam Mengelompokkan Kabupaten/Kota Berdasarkan Indikator IPM Di Jawa Tengah. *Emerging Statistics And Data Science Journal*, 3(3), 674–685. <https://doi.org/10.20885/Esds.Vol3.Iss.3.Art20>
- Handayanna, F. (2023). PENERAPAN ALGORITMA K-MEANS UNTUK KLASTERISASI PENDUDUK MISKIN DI PROVINSI BANTEN. *INTI Nusa Mandiri*, 18(1), 93–99. <https://doi.org/10.33480/Inti.V18i1.4399>
- Jaya, I. G. N. M., Nusirwan, Kusumasari, D., Susenna, A., Agustina, L., Sukma, Y. A. A., Prasetyono, H., Pangastuti, S. S., Kristiani, F., & Hermina, N. (2026). Regional Patterns Of Digital Skills Mismatch In Indonesia's Digital Economy: Insights From The Indonesia Digital Society Index. *Sustainability*, 18(2), 1077. <https://doi.org/10.3390/Su18021077>
- Jaya, I. G. N. M., Pahlevi, S. M., Susenna, A., Agustina, L., Kusumasari, D., Sukma, Y. A. A., Hemikawati, D., Rahmi, A. A., Pravitasari, A. A., & Kristiani, F. (2024). Framework For Monitoring The Spatiotemporal Distribution And Clustering Of The Digital Society Index Of Indonesia. *Sustainability*, 16(24), 11258. <https://doi.org/10.3390/Su162411258>
- Kartiasih, F., Djalal Nachrowi, N., Wisana, I. D. G. K., & Handayani, D. (2023a). Inequalities Of Indonesia's Regional Digital Development And Its Association With Socioeconomic Characteristics: A Spatial And Multivariate Analysis. *Information Technology For Development*, 29(2–3), 299–328. <https://doi.org/10.1080/02681102.2022.2110556>
- Kartiasih, F., Djalal Nachrowi, N., Wisana, I. D. G. K., & Handayani, D. (2023b). Inequalities Of Indonesia's Regional Digital Development And Its Association With Socioeconomic Characteristics: A Spatial And Multivariate Analysis. *Information Technology For Development*, 29(2–3), 299–328. <https://doi.org/10.1080/02681102.2022.2110556>
- Matdoan, M. Y., Igo, L., Rumeon, R., Fadhillah, R., & Laamena, N. S. (2024). Penerapan Algoritma K-Means Untuk Klusterisasi Kabupaten/Kota Berdasarkan Tingkat Kemiskinan Di Kepulauan Maluku Dan Papua. *Jurnal Sains Matematika Dan Statistika*, 10(1), 1–9.
- Oh, W., Acquisti, A., & Sia, C. (2018). ICT Challenges And Opportunities In Building A “Bright Society.” *Journal Of The Association For Information Systems*, 19(2), 58–62. <https://doi.org/10.17705/Jais1.00483>
- Roztock, N., Soja, P., & Weistroffer, H. R. (2019). The Role Of Information And Communication Technologies In Socioeconomic Development: Towards A Multi-Dimensional Framework. *Information Technology For Development*, 25(2), 171–183. <https://doi.org/10.1080/02681102.2019.1596654>
- Simbolon, P. S., Tarigan, E. D., & Herianto, T. J. (2025). The Application Of The K-Means Clustering Method In Grouping The Number Of Poor People Per Regency/City In Papua Province In 2023. *Journal Of Mathematics Technology And Education*, 4(2), 227–238. <https://doi.org/10.32734/Jomte.V4i2.22466>
- Warschauer, M., & Matuchniak, T. (2010). New Technology And Digital Worlds: Analyzing Evidence Of Equity In Access, Use, And Outcomes. *Review Of Research In Education*, 34(1), 179–225. <https://doi.org/10.3102/0091732X09349791>
- Xia, L., Baghaie, S., & Mohammad Sajadi, S. (2024). The Digital Economy: Challenges And Opportunities In The New Era Of Technology And Electronic Communications. *Ain Shams Engineering Journal*, 15(2), 102411. <https://doi.org/10.1016/J.Asej.2023.102411>