

IDENTIFIKASI POLA TINGKAT KESENJANGAN KETUNTASAN PENDIDIKAN DI INDONESIA DENGAN MENGGUNAKAN METODE K-MEDOIDS CLUSTERING

Antika Zahrotul Kamalia^{*1}, Ismasari Nawangsih²

^{1,2} Universitas Pelita Bangsa, Bekasi

Email: ^{1*}antika.kamalia@pelitabangsa.ac.id, ²ismasari.n@pelitabangsa.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 25 Januari 2025, diterima untuk diterbitkan: 13 April 2025)

Abstrak

Pendidikan memiliki peran penting dalam pembangunan sosial dan ekonomi, namun masih ada kesenjangan ketuntasan pendidikan antar provinsi di Indonesia. Kesenjangan ini disebabkan oleh perbedaan akses fasilitas, kualitas pengajaran, dan kondisi ekonomi, yang mempengaruhi tingkat ketuntasan pendidikan. Penelitian ini menggunakan metode K-Medoids Clustering untuk mengelompokkan wilayah di Indonesia berdasarkan data ketuntasan pendidikan dari Badan Pusat Statistik Indonesia (BPS) 2018-2023 untuk sebagai pusat *Cluster* (medoids), dan dipilih karena keunggulannya dalam mengelompokkan data berdasarkan median, yang pada gilirannya membuatnya lebih tahan terhadap pengaruh data outlier. Temuan utama penelitian ini menunjukkan kesenjangan pendidikan tinggi yaitu berada di wilayah Indonesia bagian tengah dan timur yaitu seperti provinsi Nusa Tenggara Timur (NTT) dan provinsi Papua, menjadi refleksi nyata dari keterbatasan akses dan fasilitas pendidikan di wilayah tersebut. Rendahnya tingkat ketuntasan pendidikan pada jenjang SD, SMP, dan SMA turut berdampak pada minimnya angka partisipasi di pendidikan tinggi dan disparitas signifikan dalam ketuntasan pendidikan dengan kondisi seperti ini mencerminkan perlunya kebijakan berbasis data, peningkatan infrastruktur pendidikan, serta program dukungan untuk siswa dan guru guna memperbaiki kesenjangan pendidikan. Wilayah Indonesia bagian barat dengan ketuntasan pendidikan lebih tinggi, seperti Jawa, Sumatra, dan Bali, diharapkan dapat berbagi praktik terbaik untuk mendukung wilayah yang tertinggal.

Kata kunci: Pola, Kesenjangan, Pendidikan, K-Medoids

IDENTIFICATION OF PATTERNS IN EDUCATIONAL COMPLETION DISPARITIES IN INDONESIA USING THE K-MEDOIDS CLUSTERING METHOD

Abstract

Education plays a crucial role in social and economic development; however, there are still significant disparities in educational completion rates across provinces in Indonesia. These disparities are caused by differences in access to facilities, teaching quality, and economic conditions, which affect educational outcomes. This study uses the K-Medoids Clustering method to group regions in Indonesia based on educational attainment data from Statistics Indonesia (BPS) 2018-2023, using this data as the cluster centers (medoids). This method was chosen for its effectiveness in clustering data based on the median, making it more resilient to the influence of outlier data. The main findings of this study reveal significant educational disparities in central and eastern Indonesia, such as in East Nusa Tenggara (NTT) and Papua provinces, reflecting the limited access to and availability of educational facilities in these regions. Low completion rates at the elementary, junior high, and high school levels also affect participation in higher education and significant disparities in educational attainment. This underscores the need for data-driven policies, improved educational infrastructure, and support programs for students and teachers to address educational inequalities. Western regions of Indonesia with higher educational attainment, such as Java, Sumatra, and Bali, are expected to share best practices to support underdeveloped regions.

Keywords: Pattern, Disparity, Education, K-Medoids

1. PENDAHULUAN

Pendidikan merupakan faktor krusial dalam pembangunan sosial dan ekonomi suatu negara, namun di Indonesia masih terdapat kesenjangan ketuntasan pendidikan yang signifikan antara berbagai provinsi. Pendidikan merupakan suatu proses yang bertujuan untuk memaksimalkan kemampuan individu (Richia Putri et al., 2023). Dalam penelitian yang dilakukan oleh Takyi pada tahun 2021 mengenai kesenjangan pendidikan di negara Ghana menyatakan bahwa kesenjangan pendidikan sering kali dipengaruhi oleh perbedaan akses ke fasilitas pendidikan, lokasi geografis, kualitas pengajaran, dan kondisi ekonomi lokal, yang menyebabkan variasi dalam tingkat ketuntasan pendidikan (Takyi et al., 2021). Untuk memahami dan mengatasi kesenjangan ini secara efektif, diperlukan analisis yang dapat mengidentifikasi pola-pola distribusi pendidikan di berbagai wilayah. Metode *K-Medoids Clustering* menawarkan pendekatan yang efektif untuk mengelompokkan daerah berdasarkan tingkat ketuntasan pendidikan dengan menggunakan data nyata sebagai pusat *Cluster* (medoids). *K-Medoids Clustering* terpilih dengan kemampuannya yang unggul mengelompokkan data menurut median. Keunggulan ini membuat metode tersebut lebih tahan terhadap pengaruh data *outlier*, sehingga menghasilkan analisis yang lebih akurat dan stabil (M. Puja Alif Budiman & Winarso, 2024).

Dalam penelitian yang dilakukan oleh Marlina, Fernando, & Ramadhan pada tahun 2018, yang membandingkan algoritma *K-Medoids* dengan algoritma *K-Means*, dapat disimpulkan bahwa *K-Medoids* memberikan hasil validitas yang lebih baik dalam mengelompokkan data mengenai penyebaran anak cacat. Berbeda dengan algoritma *K-Means*, *K-Medoids* menunjukkan performa yang lebih superior dalam hal ketepatan dan akurasi pengelompokan, sehingga lebih efektif untuk jenis data tersebut. (Gita Aprilianu, 2022). Algoritma *K-Medoids* menawarkan keunggulan signifikan dengan mengatasi kelemahan yang sering ditemukan pada algoritma *K-Means*, terutama dalam hal sensitivitas terhadap *noise* dan *outlier*. Sementara *K-Means* cenderung terpengaruh oleh data yang tidak normal atau ekstrem, yang dapat mempengaruhi hasil clustering secara negatif, *K-Medoids* menggunakan medoids sebagai pusat klaster, yang memberikan kestabilan lebih dalam situasi di mana data mengandung *noise* atau *outlier*. Dengan demikian, *K-Medoids* dapat menghasilkan pengelompokan yang lebih akurat dan andal, bahkan ketika menghadapi data yang tidak sempurna atau terdistorsi, sehingga meningkatkan kualitas dan efektivitas proses *clustering* secara keseluruhan (Rahayu et al., 2020).

Penelitian ini bertujuan untuk mengidentifikasi pola-pola kesenjangan dalam tingkat ketuntasan pendidikan di seluruh wilayah bagian Indonesia, dengan memanfaatkan metode *K-Medoids Clustering*

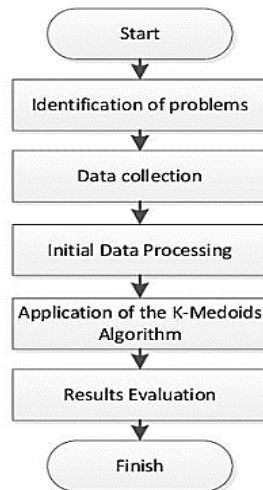
sebagai alat utama. Metode ini dipilih karena kemampuannya untuk mengelompokkan data secara efektif, memungkinkan peneliti untuk mengungkap variasi dan disparitas dalam penyelesaian pendidikan di berbagai daerah. Dengan menganalisis hasil clustering, penelitian ini juga bertujuan untuk memberikan wawasan mendalam mengenai perbedaan dalam pencapaian pendidikan, serta untuk mengidentifikasi provinsi-provinsi di Indonesia yang membutuhkan perhatian dan intervensi lebih lanjut dalam upaya meratakan kualitas pendidikan di seluruh provinsi. Oleh karena itu, penelitian ini diharapkan mampu memberikan pemahaman yang lebih mendalam tentang distribusi pendidikan di berbagai provinsi di Indonesia. Selain itu, hasil yang diperoleh dari penelitian ini diharapkan dapat memberikan kontribusi dalam merancang kebijakan pendidikan yang lebih efektif dan terarah, sehingga berperan dalam upaya mengurangi ketidakmerataan serta meningkatkan kualitas pendidikan di seluruh Indonesia.

2. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah data mining dengan pendekatan *K-Medoids*. Data mining, atau penambangan data, terdiri dari serangkaian proses yang dirancang untuk mengekstraksi, menganalisis, dan mengolah informasi berharga dari kumpulan data yang sangat besar. Tujuan utama dari proses ini adalah untuk mengidentifikasi pola, tren, atau hubungan yang mungkin tersembunyi atau belum terdeteksi sebelumnya. Dengan cara ini, data mining berfungsi untuk mengungkap wawasan baru dan menyediakan pemahaman yang lebih mendalam mengenai data, yang pada gilirannya dapat digunakan untuk membuat keputusan yang lebih baik serta merancang strategi yang lebih efektif (M. Puja Alif Budiman & Winarso, 2024).

Pendekatan mendasar dalam menggali inti dari data dan mengekstraksi informasi yang berharga, yang sering kali tersembunyi, melibatkan penggunaan berbagai teknik dan alat analitis. Proses ini bertujuan untuk mengungkap keterkaitan dan pola-pola yang mungkin tidak terlihat secara jelas. Dengan menerapkan berbagai metode analisis, kita dapat menerangi informasi yang terperangkap dalam kompleksitas data dan menemukan hubungan-hubungan penting yang sebelumnya mungkin terabaikan. Pendekatan ini membantu dalam mengidentifikasi wawasan yang signifikan dan memberikan pemahaman yang lebih mendalam tentang data yang ada (Rahayu et al., 2020). Data mining terdiri dari serangkaian tahapan yang dirancang untuk mengungkap nilai tambah dari kumpulan data, dengan tujuan menghasilkan pengetahuan baru yang sebelumnya tidak dapat diperoleh secara manual. Proses ini melibatkan ekstraksi kecenderungan dan pola-pola data yang tersembunyi, dan selanjutnya mengubah hasil temuan

tersebut menjadi informasi yang akurat dan mudah dipahami. Dengan melalui proses data mining, kita dapat mengidentifikasi pola-pola yang signifikan dan transformasi hasil analisis menjadi wawasan yang berguna, yang pada gilirannya mempermudah pemahaman dan aplikasi informasi tersebut dalam konteks yang relevan (Zahrotul Kamalia et al., 2023). Berikut diagram alir dalam penelitian ini:



Gambar 1. Diagram Alir Penelitian

Alat yang digunakan dalam pengolahan data termasuk *Rapidminer*, yang merupakan salah satu solusi data mining yang paling populer. *Rapidminer* berfungsi sebagai alat yang sangat efektif karena menyediakan beragam metode analitis dan teknik yang diperlukan. Alat ini mencakup berbagai fungsi, mulai dari evaluasi statistik sederhana seperti analisis korelasi, hingga teknik yang lebih kompleks seperti regresi, klasifikasi, dan pengelompokan. Selain itu, *Rapidminer* juga mendukung pengurangan dimensi dan optimasi parameter, memungkinkan pengguna untuk mengolah dan menganalisis data secara komprehensif. Dengan kemampuannya yang luas, *Rapidminer* mempermudah proses pengolahan data dan membantu dalam mengungkap wawasan berharga dari kumpulan data yang besar dan kompleks. (Junaedi et al., 2022).

Rapidminer juga memungkinkan hasil analisis ditampilkan secara visual melalui berbagai grafik, yang memudahkan pemahaman dan interpretasi data. Fitur visualisasi ini membuat *Rapidminer* menjadi salah satu pilihan utama dalam proses ekstraksi data, karena tidak hanya menyediakan analisis mendalam tetapi juga menyajikan hasilnya dalam format yang jelas dan mudah diinterpretasikan. Dengan kemampuannya untuk menggabungkan analisis yang kuat dengan visualisasi yang intuitif, *Rapidminer* memfasilitasi pemahaman yang lebih baik tentang pola-pola data dan keterkaitan yang teridentifikasi, menjadikannya alat yang sangat berharga dalam aplikasi metode-metode data mining. (Chandra et al., 2021).

2.1 Pengumpulan Data

Sumber data untuk penelitian ini diperoleh dari Badan Pusat Statistik (BPS) Indonesia, mencakup periode enam tahun dari 2018 hingga 2023. Data tersebut melibatkan seluruh 34 provinsi di Indonesia sebagai negara kepulauan, terbagi menjadi tiga wilayah geografis utama yang mencakup Indonesia bagian barat, tengah, dan timur. Wilayah barat meliputi Aceh, Sumatra Utara, Sumatra Barat, Riau, Jambi, Sumatra Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, dan Banten. Wilayah tengah mencakup 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, Sulawesi Barat, dan Gorontalo. Sedangkan wilayah timur mencakup Maluku, Maluku Utara, Papua Barat, dan Papua.

Dalam penelitian ini, data yang digunakan meliputi tingkat penyelesaian pendidikan di jenjang SD, SMP, dan SMA selama periode 2018 hingga 2023. Data ini memberikan gambaran menyeluruh mengenai ketuntasan pendidikan di seluruh provinsi di Indonesia, dan memungkinkan analisis mendalam mengenai variabilitas serta tren dalam pencapaian pendidikan di berbagai wilayah. Dengan demikian, penelitian ini bertujuan untuk memahami dan mengidentifikasi tingkat pola kesenjangan pendidikan di seluruh wilayah bagian Indonesia.

2.2 Clustering K-Medoids

Algoritma K-Medoids, atau Partitioning Around Medoids (PAM), adalah metode clustering yang membagi objek menjadi beberapa cluster dengan memilih medoids sebagai perwakilan paling representatif di setiap cluster. Algoritma ini membantu mengorganisasi data secara efektif, mengidentifikasi pola tersembunyi, dan menghasilkan cluster yang mencerminkan struktur data secara akurat. Dengan pertumbuhan data yang pesat, K-Medoids menjadi solusi penting untuk mengelola data besar dan kompleks, mempermudah analisis serta pemahaman informasi (Harahap, 2021).

Algoritma K-Medoids memanfaatkan objek dalam kumpulan data untuk berfungsi sebagai pusat representatif dari masing-masing cluster. Dalam metode ini, objek yang dipilih untuk mewakili sebuah cluster dikenal sebagai medoids. Medoids berperan sebagai representasi yang paling representatif dari cluster tersebut, sehingga memudahkan dalam penentuan pusat dan pembentukan kelompok data yang lebih homogen. Dengan menggunakan medoids sebagai pusat cluster, algoritma ini dapat mengelompokkan data secara lebih efektif dan akurat, serta memberikan pemahaman yang lebih jelas tentang struktur data yang dianalisis. Dalam

Tabel 1. Ketuntasan Sekolah Dasar (SD)

Provinsi	2018	2019	2020	2021	2022	2023
Aceh	96.75	98.6	98.33	99.44	99.45	99.08
Sumatera Utara	95.76	97.12	96.48	98.57	98.74	98.75
Sumatera Barat	90.54	91.25	93.26	95.29	97.87	95.81
Riau	95.39	95.11	95.18	96.91	98.2	98.09
Jambi	94.62	96.99	95.34	98.54	97.62	97.76
.						
.						
Maluku	94.49	94.95	97.47	98.5	98.98	98.69
Maluku Utara	95.65	94.66	95.59	96.97	97.72	98.3
Papua Barat	86.46	88.46	89.25	91.81	93.94	92.69
Papua	71.83	73.07	78.4	78.43	81.99	80.09

Tabel 2. Ketuntasan Sekolah Menengah Pertama (SMP)

Provinsi	2018	2019	2020	2021	2022	2023
Aceh	87.15	89.5	90.92	93.43	97.63	94.55
Sumatera Utara	87.97	88.26	91.63	91.35	92.84	94.35
Sumatera Barat	82.59	84.07	87.12	89.49	88.83	90.65
Riau	83.45	84.97	86.84	87.11	88.53	90.52
Jambi	83.29	84.4	86.31	89	86.57	89.35
.						
.						
Maluku	86.38	88.43	88.04	93.08	93.65	93.9
Maluku Utara	85.39	85.07	87.41	92.93	94.92	93.46
Papua Barat	81.19	80.91	83.47	85.18	87.03	88.63
Papua	57.19	59.31	65.75	66.06	66.16	67.12

Tabel 3. Ketuntasan Sekolah Menengah Atas (SMA)

Provinsi	2018	2019	2020	2021	2022	2023
Aceh	70.68	69.96	70.07	74.36	70.67	74.46
Sumatera Utara	68.34	65.21	70.39	72.81	77.16	74.43
Sumatera Barat	65.34	60.32	67.11	70.06	65.96	68.64
Riau	63.71	58.78	66.62	68.94	66.91	67.79
Jambi	66.06	56.87	63.66	64.51	65.85	66.62
.						
.						
Maluku	66.42	67.82	70.55	68.12	72.08	75.01
Maluku Utara	60.07	59.13	66.52	66.95	67.1	64.61
Papua Barat	60.47	50.95	61.49	59.08	57.07	59.99
Papua	29.56	27.44	30.92	32.95	39.01	39.5

Penelitian ini, *Cluster* akan dibagi menjadi tiga kategori: K=0 (C1) *Cluster* tinggi, (C2) *Cluster* sedang, dan K=2 (C3) *Cluster* rendah. *Cluster* tinggi

menunjukkan provinsi dengan tingkat kesenjangan ketuntasan pendidikan yang tinggi, *Cluster* sedang menunjukkan nilai kesenjangan ketuntasan pendidikan yang sedang, dan *Cluster* rendah menunjukkan nilai kesenjangan ketuntasan pendidikan yang rendah. Berikut merupakan langkah-langkah dalam perhitungan algoritma K-Medoids diantaranya (Hardiyanti et al., 2019):

1. Inisialisasikan pusat cluster sesuai dengan jumlah cluster yang diinginkan (k).
2. Alokasikan setiap data (objek) ke cluster terdekat dengan menggunakan ukuran jarak Euclidean Distance, yang dihitung dengan persamaan berikut:

$$d(x_{ij}, c_{kj}) = \sqrt{\sum_{j=i}^p \sum_{i=1}^n (x_{ij} - c_{kj})^2}$$

Dimana:

D (x_{ij}, c_{kj}): Jarak Euclidean Distance antara observasi ke-i pada variabel ke-j dengan pusat cluster ke-k pada variabel ke-j

X_{ij}: nilai objek pada observasi ke-i untuk variabel ke-j

c_{kj}: nilai pusat cluster ke-k pada variabel ke-j

p: jumlah variabel yang diamati

n: jumlah observasi yang diamati

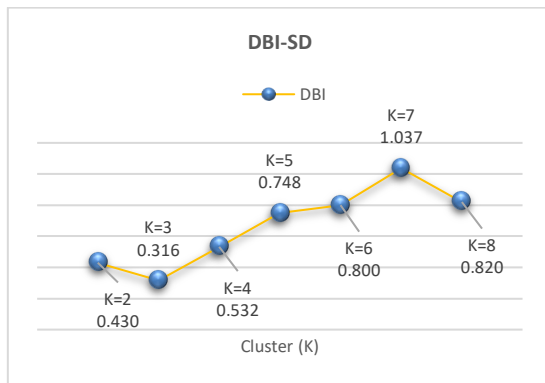
3. Pilih secara acak objek dari setiap cluster untuk dijadikan kandidat medoid baru.
4. Hitung jarak antara setiap objek di dalam cluster dengan kandidat medoid yang baru.
5. Hitung total deviasi (S) dengan membandingkan total jarak baru dengan total jarak sebelumnya. Jika nilai S kurang dari 0, lakukan pertukaran objek dengan data cluster untuk membentuk sekumpulan k objek baru sebagai medoid.
6. Ulangi langkah 3 hingga 5 sampai tidak ada lagi perubahan pada medoid, sehingga diperoleh cluster yang stabil beserta anggotanya masing-masing.

2.3 Evaluasi Davies Bouldin Indeks (DBI)

DBI merupakan teknik evaluasi internal yang digunakan untuk menilai kualitas kluster dalam pengelompokan data. Metode ini mengukur efektivitas proses pengelompokan dengan mengevaluasi baik jumlah maupun kualitas kluster yang terbentuk dari data yang telah dikumpulkan. Setelah data dipersiapkan secara menyeluruh, langkah selanjutnya adalah memproses data tersebut menggunakan algoritma K-Medoids. Dalam konteks penelitian ini, jumlah cluster ditentukan bervariasi antara 2 hingga 8, dan diterapkan pada berbagai jenjang pendidikan, termasuk Sekolah Dasar (SD), Sekolah Menengah Pertama (SMP), dan Sekolah Menengah Atas (SMA). Dengan pendekatan ini,

diharapkan dapat diperoleh pemahaman yang lebih baik mengenai struktur data di setiap jenjang pendidikan dan efektivitas pengelompokan yang dilakukan.

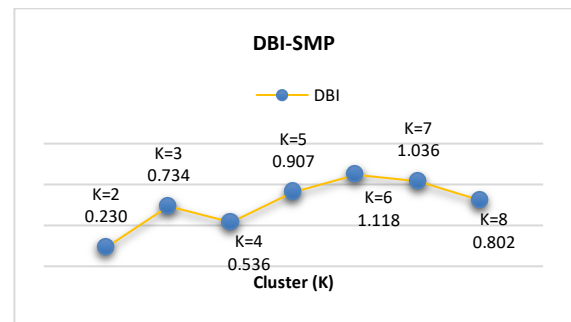
Jumlah cluster yang digunakan dalam penelitian ini ditentukan berdasarkan hasil evaluasi yang bertujuan untuk menetapkan jumlah cluster optimal pada setiap jenjang pendidikan. Evaluasi ini dilakukan dengan mempertimbangkan nilai DBI, di mana pengelompokan yang lebih baik diidentifikasi melalui nilai DBI yang lebih rendah. Dengan kata lain, penentuan jumlah cluster didasarkan pada upaya untuk mencapai nilai DBI terendah, yang mencerminkan pengelompokan yang lebih efektif dan representatif untuk masing-masing jenjang pendidikan yang diteliti. Pendekatan ini memastikan bahwa hasil pengelompokan tidak hanya sesuai dengan data yang ada, tetapi juga memberikan kualitas pengelompokan yang terbaik untuk analisis lebih lanjut (Hidayat & Rina, 2022). Dalam penelitian ini, alat analisis data yang digunakan adalah RapidMiner, yang berfungsi sebagai perangkat lunak utama untuk melaksanakan proses analisis. RapidMiner memfasilitasi penggunaan berbagai metode analisis, termasuk penerapan algoritma K-Medoids dalam proses clustering. Hasil dari evaluasi Indeks Davies-Bouldin (DBI) yang diperoleh melalui metode clustering dengan menggunakan algoritma K-Medoids pada aplikasi RapidMiner dapat dilihat pada gambar berikut. Gambar 2 tersebut menunjukkan secara visual bagaimana DBI diukur dan memberikan gambaran mengenai efektivitas pengelompokan yang dilakukan dalam penelitian ini



Gambar 2. Evaluasi Cluster Jenjang Pendidikan SD

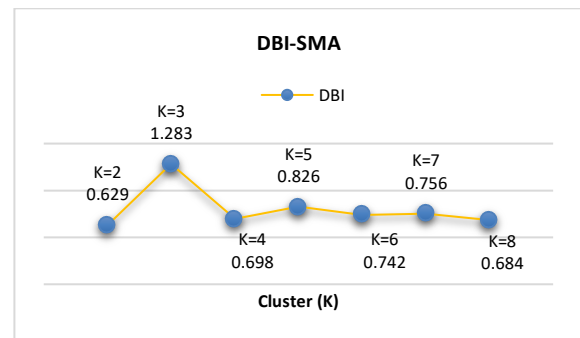
Berdasarkan gambar 2 di atas, yang menunjukkan hasil evaluasi cluster untuk jenjang pendidikan Sekolah Dasar (SD) menggunakan algoritma K-Medoids, evaluasi dilakukan dengan menentukan nilai Indeks Davies-Bouldin (DBI) untuk masing-masing klaster. Hasil rekapitulasi DBI, yang ditampilkan untuk nilai k yang bervariasi dari 2 hingga 8, menunjukkan bahwa nilai k yang optimal adalah k=3. Hal ini disebabkan oleh k=3 yang menghasilkan nilai DBI terendah, yaitu 0.316. Oleh karena itu, dalam penelitian ini, jumlah cluster yang dipilih untuk jenjang pendidikan SD adalah sebanyak 3 cluster. Pendekatan ini memastikan bahwa

pengelompokan yang dilakukan memberikan hasil yang paling efisien dan representatif berdasarkan evaluasi DBI yang dilakukan.



Gambar 3. Evaluasi Cluster Jenjang Pendidikan SMP

Berdasarkan gambar 3 di atas, yang menyajikan hasil evaluasi cluster untuk jenjang pendidikan Sekolah Menengah Pertama (SMP) menggunakan algoritma K-Medoids, analisis dilakukan dengan menghitung nilai Indeks Davies-Bouldin (DBI) untuk setiap klaster. Hasil rekapitulasi DBI, yang diperoleh untuk nilai k yang bervariasi antara 2 hingga 8, menunjukkan bahwa nilai k yang paling optimal adalah k=2. Hal ini disebabkan oleh k=2 yang menghasilkan nilai DBI terendah, yaitu 0.230. Oleh karena itu, dalam penelitian ini, jumlah cluster yang dipilih untuk jenjang pendidikan SMP adalah sebanyak 2 cluster. Keputusan ini diambil untuk memastikan bahwa pengelompokan yang dilakukan memberikan hasil yang paling efisien dan representatif berdasarkan evaluasi DBI yang telah dilakukan.



Gambar 3. Evaluasi Cluster Jenjang Pendidikan SMA

Berdasarkan gambar 3 di atas, yang menampilkan hasil evaluasi cluster untuk jenjang pendidikan Sekolah Menengah Atas (SMA) dengan menggunakan algoritma K-Medoids, evaluasi dilakukan dengan menentukan nilai Indeks Davies-Bouldin (DBI) untuk setiap klaster. Hasil rekapitulasi DBI, yang mencakup nilai k dari 2 hingga 8, menunjukkan bahwa nilai k yang paling optimal adalah k=2, karena pada percobaan ini diperoleh nilai DBI terendah, yaitu 0.629. Oleh karena itu, dalam penelitian ini, jumlah cluster yang dipilih untuk jenjang pendidikan SMA adalah sebanyak 2 cluster. Keputusan ini diambil untuk memastikan bahwa pengelompokan yang dilakukan adalah yang paling

efisien dan representatif, berdasarkan evaluasi DBI yang telah dilakukan.

3. HASIL DAN PEMBAHASAN

Atribut yang digunakan dalam pengolahan data terdiri dari tingkat pendidikan dan provinsi, sedangkan atribut kelas data yang dipertimbangkan adalah persentase ketuntasan pendidikan. Dataset yang relevan disimpan dalam format CSV dan kemudian diimpor ke dalam aplikasi analisis. Setelah proses impor selesai, data tersebut dimasukkan ke dalam aplikasi dan dilengkapi dengan menambahkan operator data serta operator K-Medoids di jendela proses. Operator data berfungsi untuk mengelola data, sedangkan operator K-Medoids diterapkan untuk menjalankan algoritma K-Medoids pada dataset yang telah disiapkan.

Pada tahap konfigurasi operator K-Medoids, pengaturan jumlah cluster (k) disesuaikan dengan hasil evaluasi Indeks Davies-Bouldin (DBI) yang telah dilakukan sebelumnya. Selain itu, pengukuran jarak antara data dan centroid dilakukan menggunakan Euclidean Distance, yang memungkinkan penentuan jarak yang tepat untuk masing-masing data terhadap pusat cluster. Pendekatan ini memastikan bahwa proses clustering berjalan dengan efisien dan akurat, berdasarkan parameter yang telah ditetapkan dan analisis yang telah dilakukan.

3.1 Identifikasi pola pada jenjang pendidikan SD

Hasil dari proses pengolahan data ketuntasan pendidikan jenjang SD menggunakan algoritma K-Medoids dengan jumlah kluster yaitu 3 menggunakan alat *Rapidminer* pada tabel 4 dibawah ini:

Tabel 4. Cluster Model SD

Cluster	Banyak Data	Data dalam Cluster (%)
C1	1	3%
C2	4	12%
C3	29	85%
Total	34	100%

Pada Tabel 4, hasil output menunjukkan bahwa data kluster untuk jenjang pendidikan Sekolah Dasar (SD) terbagi menjadi tiga kluster. Kluster pertama, C1, terdiri dari satu provinsi yaitu Papua yang merupakan Indonesia bagian timur. Kluster kedua, C2, mencakup empat provinsi, yakni Nusa Tenggara Timur, Kalimantan Barat, Gorontalo, dan Papua Barat yang merupakan Indonesia bagian tengah dan timur. Sementara itu, kluster ketiga, C3, mencakup 29 provinsi yang meliputi Aceh, Sumatra Utara, Sumatra Barat, Riau, Jambi, Sumatra Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara,

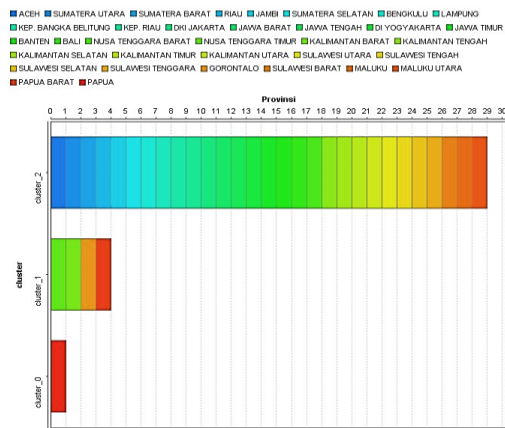
Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Sulawesi Barat, Maluku, dan Maluku Utara yang merupakan rata-rata didominasi Indonesia bagian barat dan sebagianya lagi bagian tengah. Pembagian kluster ini memberikan gambaran menyeluruh mengenai distribusi provinsi di Indonesia berdasarkan hasil clustering yang dilakukan untuk jenjang pendidikan SD.

Tabel 5. Nilai centroid SD

Attribute	C1	C2	C3
2018	71.830	86.460	95.650
2019	73.070	88.460	94.660
2020	78.400	89.250	95.590
2021	78.430	91.810	96.970
2022	81.990	93.940	97.720
2023	80.090	92.690	98.300

Pada Tabel 5, ditunjukkan bagaimana penentuan nilai kluster dikategorikan menjadi tinggi, sedang, dan rendah berdasarkan nilai centroid. Tabel ini menguraikan hasil akhir dari pemetaan jenjang pendidikan Sekolah Dasar (SD) yang dilakukan dengan menggunakan nilai centroid. Proses ini melibatkan analisis centroid untuk setiap kluster, yang digunakan sebagai dasar untuk mengklasifikasikan kluster-kluster tersebut ke dalam kategori tingkat ketuntasan yang berbeda. Dengan cara ini, Tabel 5 memberikan gambaran jelas mengenai bagaimana setiap kluster dikelompokkan sesuai dengan nilai centroidnya, mencerminkan distribusi dan variasi ketuntasan pendidikan di jenjang SD.

Nilai centroid menggambarkan rata-rata ketuntasan pendidikan untuk setiap *Cluster* dalam rentang waktu yang ditentukan. Nilai centroid C1 menunjukkan pola yang relatif lebih rendah dibandingkan dengan C2 dan C3. Terdapat peningkatan yang konsisten dari 71.830 pada tahun 2018 menjadi 80.090 pada tahun 2023. Ini menunjukkan bahwa meskipun ada peningkatan, *Cluster* ini cenderung memiliki ketuntasan pendidikan yang lebih rendah secara keseluruhan. C2 memiliki nilai centroid yang lebih tinggi dibandingkan dengan C1 tetapi lebih rendah dibandingkan dengan C3. Terdapat juga peningkatan nilai dari 86.460 pada tahun 2018 menjadi 92.690 pada tahun 2023. Pola ini menunjukkan bahwa ketuntasan pendidikan di *Cluster* ini lebih baik daripada C1 dan mengalami pertumbuhan yang stabil. C3 memiliki nilai centroid tertinggi dibandingkan dengan *Cluster* lainnya, menunjukkan ketuntasan pendidikan yang paling tinggi secara keseluruhan. Nilai centroid C3 meningkat dari 95.650 pada tahun 2018 menjadi 98.300 pada tahun 2023. Ini menunjukkan bahwa *Cluster* ini memiliki ketuntasan pendidikan terbaik dan terus mengalami peningkatan.



Gambar 4 Visualisasi Clustering SD

Visualisasi data pada gambar 4 menunjukkan bahwa C1 yang hanya mencakup Papua yang merupakan Indonesia bagian timur, memiliki ketuntasan pendidikan SD yang jauh lebih rendah dibandingkan provinsi lain sehingga termasuk kategori kesenjangan yang tinggi dan mencerminkan tantangan besar seperti kesulitan geografis dan keterbatasan infrastruktur. C2 yang meliputi Nusa Tenggara Timur, Kalimantan Barat, Gorontalo, dan Papua Barat yang merupakan Indonesia bagian tengah dan timur, juga menghadapi tantangan aksesibilitas dan infrastruktur, meskipun ketuntasannya lebih baik dari Papua. Sebaliknya, C3 yang terdiri dari 29 provinsi di pulau Jawa, Sumatra, Kalimantan, dan Sulawesi yang merupakan rata-rata Indonesia bagian barat dan sebagian bagian tengah, menunjukkan kemajuan signifikan dalam ketuntasan pendidikan, berkat infrastruktur dan akses yang lebih baik. Sehingga dalam jenjang pendidikan SD terdapat kesenjangan yang jelas dalam ketuntasan pendidikan antara Cluster dengan jumlah provinsi yang lebih kecil, sedang dan yang lebih besar, mengindikasikan perlunya penyesuaian kebijakan pendidikan untuk menangani disparitas geografis dan meningkatkan kualitas pendidikan di wilayah Indonesia yang kurang berkembang.

Beberapa faktor penyebab ketimpangan pendidikan di wilayah Indonesia timur yaitu Papua Barat meliputi sejumlah isu krusial. Salah satunya adalah rendahnya tingkat ketersediaan sekolah, khususnya untuk jenjang pendidikan selain sekolah dasar, di daerah-daerah yang baru dimekarkan dan di wilayah-wilayah terpencil. Selain itu, biaya pendidikan yang dianggap terlalu tinggi seringkali menjadi hambatan signifikan bagi banyak keluarga, mengakibatkan kesulitan dalam melanjutkan proses pendidikan. Untuk mengatasi masalah ini, penting dilakukan sosialisasi yang intensif kepada masyarakat mengenai pentingnya anak-anak untuk menyelesaikan pendidikan dasar wajib selama sembilan tahun. Upaya sosialisasi ini diharapkan dapat meningkatkan kesadaran dan partisipasi masyarakat dalam memastikan bahwa semua anak mendapatkan akses dan kesempatan untuk

menyelesaikan pendidikan dasar yang wajib (Nisa & Samputra, 2020).

Salah satu faktor utama yang mempengaruhi ketidakmerataan kualitas pendidikan antara kawasan urban dan pedalaman adalah ketidakcukupan sarana dan prasarana yang tersedia. Contohnya yang sering dijumpai adalah akses internet. Di kawasan kota, akses internet sangat mudah didapatkan; hampir setiap siswa Sekolah Dasar (SD) memiliki smartphone mereka sendiri, yang dianggap sebagai hal yang umum. Bahkan, banyak orang tua di kota yang memasang jaringan Wi-Fi dengan kapasitas besar di rumah mereka. Sebaliknya, di desa-desa atau daerah pedalaman, memiliki handphone sudah dianggap luar biasa di mata anak-anak lainnya. Selain itu, masih ada wilayah-wilayah terpencil yang belum terjangkau oleh jaringan seluler. Kondisi ini berdampak signifikan pada perkembangan kualitas pendidikan di Indonesia, mengingat hampir semua informasi dapat diakses dengan cepat melalui koneksi internet. Perbedaan ini menggambarkan betapa besar pengaruh sarana dan prasarana terhadap perbedaan kualitas pendidikan antara kota dan desa atau pedalaman. (Kogoya et al., 2023).

3.2 Identifikasi pola pada jenjang pendidikan SMP

Hasil dari proses pengolahan data ketuntasan pendidikan jenjang SMP dengan penggunaan algoritma K-Medoids total kluster yaitu 3 menggunakan alat *Rapidminer* pada tabel 6 dibawah ini:

Tabel 6. Cluster Model SMP

Cluster	Banyak Data	Data dalam Cluster (%)
C1	1	3%
C2	33	97%
Total	34	100%

Pada Tabel 6, hasil output menunjukkan pembagian kluster untuk jenjang pendidikan Sekolah Menengah Pertama (SMP) yang terbagi menjadi dua kluster. Kluster pertama, C1, mencakup satu provinsi, yaitu Papua yang merupakan Indonesia bagian timur. Sementara itu, kluster kedua, C2, terdiri dari 33 provinsi, yang meliputi Aceh, Sumatra Utara, Sumatra Barat, Riau, Jambi, Sumatra Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan 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, Sulawesi Barat, Gorontalo, Maluku, Maluku Utara, dan Papua Barat yang merupakan gabungan dari Indonesia bagian barat tengah dan timur. Pembagian ini memberikan gambaran menyeluruh mengenai distribusi provinsi

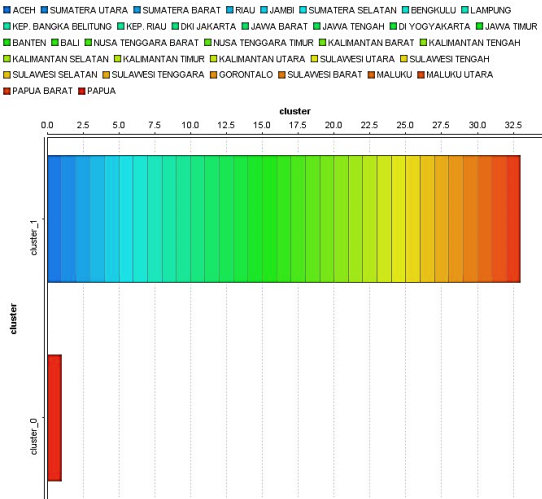
di Indonesia untuk jenjang SMP berdasarkan hasil clustering yang dilakukan.

Tabel 7. Nilai centroid SMP

Attibute	C1	C2
2018	57.190	81.190
2019	59.310	80.910
2020	65.750	83.470
2021	66.060	85.180
2022	66.160	87.030
2023	67.120	88.630

Pada Tabel 7, ditampilkan penentuan nilai klaster yang dikategorikan sebagai tinggi, sedang, dan rendah berdasarkan analisis nilai centroid. Tabel ini menggambarkan hasil akhir dari pemetaan jenjang pendidikan Sekolah Menengah Pertama (SMP), yang dilakukan dengan menggunakan nilai centroid sebagai acuan. Proses ini melibatkan evaluasi nilai centroid untuk setiap klaster, yang selanjutnya digunakan untuk mengklasifikasikan klaster-klaster tersebut ke dalam kategori tingkat ketuntasan pendidikan yang berbeda. C1 menunjukkan nilai centroid yang lebih rendah dibandingkan dengan C2 pada setiap tahun. Nilai centroid C1 meningkat dari 57.190 pada tahun 2018 dengan 67.120 pada tahun 2023. Meskipun terdapat peningkatan yang konsisten dalam ketuntasan pendidikan SMP di C1, nilai ini tetap lebih rendah dibandingkan dengan C2. Peningkatan ini menunjukkan adanya kemajuan, tetapi ketuntasan pendidikan di *Cluster* ini masih berada di bawah rata-rata ketuntasan pendidikan nasional. C2 memiliki nilai centroid yang lebih tinggi dibandingkan dengan C1. Nilai centroid C2 meningkat dari 81.190 pada tahun 2018 menjadi 88.630 pada tahun 2023. Ketuntasan pendidikan SMP di C2 menunjukkan tren yang stabil dan positif dengan nilai yang lebih tinggi setiap tahun. Ini mengindikasikan bahwa provinsi dalam *Cluster* ini mungkin mempunyai akses baik ke sumber daya pendidikan, infrastruktur yang lebih baik, atau kebijakan pendidikan yang lebih efektif.

Visualisasi data di gambar 5 menunjukkan bahwa Papua yang merupakan Indonesia bagian timur, yang termasuk dalam *Cluster* C1, memiliki kesenjangan ketuntasan pendidikan SMP yang jauh lebih rendah daripada provinsi-provinsi di wilayah Indonesia dibagian lainnya, yang tergabung dalam C2. Ketuntasan pendidikan di Papua dipengaruhi oleh faktor-faktor seperti kondisi geografis yang sulit dan akses terbatas ke fasilitas pendidikan, yang menyebabkan keterlambatan dalam pengembangan pendidikan. Sebaliknya, *Cluster* C2 yang mencakup wilayah gabungan Indonesia bagian barat, tengah dan timur lebih maju seperti pulau Jawa, Sumatra, Kalimantan, Sulawesi, dan beberapa provinsi di Nusa Tenggara serta Maluku, menunjukkan ketuntasan pendidikan yang lebih baik berkat infrastruktur dan dukungan yang lebih baik.



Gambar 5. Visualisasi *Clustering* SMP

Perbedaan ketuntasan pendidikan antara Papua dan provinsi lain mencerminkan disparitas geografis yang signifikan, yang menunjukkan perlunya program intervensi yang lebih intensif untuk Papua dan penguatan kebijakan pendidikan di wilayah yang sudah lebih maju. Pelatihan dan peningkatan kualitas sumber daya manusia (SDM) di kalangan tenaga pendidik merupakan langkah awal yang sangat krusial untuk memastikan pemerataan kualitas pendidikan di seluruh wilayah. Dengan memiliki guru-guru yang berkualitas, sekolah-sekolah yang berada di daerah terpencil sekalipun akan mampu menghasilkan hasil pendidikan yang baik dan berkualitas. Peningkatan kompetensi tenaga pendidik ini akan berkontribusi secara signifikan dalam meningkatkan standar pendidikan secara keseluruhan, sehingga menciptakan kesempatan yang sama bagi semua siswa, tidak peduli seberapa terpencil lokasi mereka. Dengan demikian, investasi dalam pelatihan dan pengembangan SDM pendidikan akan berdampak positif pada semua tingkat pendidikan dan memastikan bahwa setiap anak memiliki akses kepada pendidikan yang berkualitas.(Maulana, 2022).

3.3 Identifikasi pola pada jenjang pendidikan SMA

Hasil dari proses pengolahan data ketuntasan pendidikan jenjang SMA menggunakan algoritma K-Medoids dengan jumlah kluster yaitu 3 menggunakan alat *Rapidminer* pada tabel 5 dibawah ini.

Tabel 8. *Cluster* Model SMA

Cluster	Banyak Data	Data dalam Cluster (%)
C1	2	6%
C2	32	94%
Total	34	100%

Pada Tabel 8, hasil output menunjukkan pembagian kluster untuk jenjang pendidikan Sekolah Menengah Atas (SMA) yang terbagi menjadi dua kluster. Kluster C2 mencakup 32 provinsi, yaitu

Aceh, Sumatra Utara, Sumatra Barat, Riau, Jambi, Sumatra Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Sulawesi Barat, Gorontalo, Maluku, Maluku Utara, dan Papua Barat. Sementara itu, kluster C1 terdiri dari dua provinsi dari Indonesia bagian tengah dan timur, yakni Nusa Tenggara Timur dan Papua. Pembagian kluster ini memberikan gambaran yang jelas tentang distribusi provinsi di Indonesia pada jenjang pendidikan SMA, berdasarkan hasil clustering yang dilakukan.

Tabel 8. Nilai centroid SMA

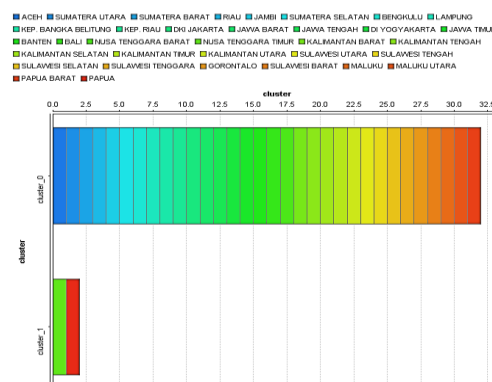
Attribute	C1	C2
2018	60.470	29.560
2019	50.590	27.440
2020	61.490	30.920
2021	59.080	32.950
2022	57.070	39.010
2023	59.990	39.500

Pada Tabel 5, ditampilkan penentuan nilai klaster yang dikategorikan sebagai tinggi, sedang, dan rendah berdasarkan analisis nilai centroid dari pemetaan jenjang pendidikan Sekolah Menengah Atas (SMA). Tabel ini menunjukkan bahwa nilai ketuntasan pendidikan SMA di klaster C1 menunjukkan tren yang sedikit fluktuatif, namun nilai rata-ratanya tetap konsisten lebih tinggi dari tahun 2018 hingga 2023. Hal ini mengindikasikan bahwa provinsi-provinsi yang tergabung dalam klaster ini mungkin memiliki akses yang lebih baik ke sumber daya pendidikan dan fasilitas yang lebih memadai dibandingkan dengan provinsi-provinsi yang termasuk dalam klaster C2. Perbedaan ini menunjukkan adanya disparitas signifikan dalam kualitas pendidikan antara kedua klaster, yang terlihat jelas dari analisis nilai centroid yang dilakukan.

Berdasarkan visualisasi pada gambar 6 menggambarkan bahwa menunjukkan adanya disparitas signifikan dalam kesenjangan ketuntasan pendidikan SMA antara provinsi di C1 dan provinsi di C2. Perbedaan ini mencerminkan ketidakmerataan atau kesenjangan dalam kualitas pendidikan di Indonesia, di mana sebagian besar provinsi di wilayah Indonesia bagian barat memiliki tingkat ketuntasan yang baik, sementara provinsi di wilayah Indonesia bagian tengah dan timur seperti Nusa Tenggara Timur dan Papua menunjukkan kebutuhan mendesak untuk perbaikan. Hasil penelitian ini menganalisis perlunya kebijakan pendidikan yang lebih terfokus untuk meningkatkan kualitas pendidikan di wilayah Indonesia yang mencakup provinsi-provinsi dengan ketuntasan yang rendah.

Secara geografis, provinsi Nusa Tenggara Timur memiliki karakteristik yang mencakup baik wilayah perkotaan maupun pedalaman. Penggunaan istilah "pedalaman" di sini tidak dimaksudkan untuk

merendahkan atau mendiskreditkan suatu daerah, melainkan untuk menekankan posisi geografisnya yang memang cukup jauh dari pusat kota.



Gambar 6 Visualisasi Clustering SMA

Wilayah pedalaman Indonesia bagian tengah di Nusa Tenggara Timur membutuhkan perhatian khusus dalam upaya pemerataan pembangunan, terutama karena adanya keterbatasan dalam hal transportasi, infrastruktur, serta kebutuhan akan tenaga pengajar yang bersedia dan siap untuk mengabdikan diri di daerah-daerah terpencil tersebut. Oleh karena itu, pengembangan dan peningkatan fasilitas serta dukungan bagi tenaga pendidik di daerah-daerah ini sangat penting untuk menciptakan keseimbangan dalam akses dan kualitas pendidikan (Metarum, 2021).

Anak-anak di wilayah Indonesia bagian timur di provinsi Papua menghadapi serangkaian tantangan yang sangat khusus, termasuk aksesibilitas yang terbatas, infrastruktur yang belum memadai, serta kekurangan tenaga pengajar yang berkualifikasi. Selain itu, terdapat ketidakmerataan dalam kualitas pendidikan yang mereka terima. Situasi ini memerlukan perhatian mendalam dan upaya yang serius dari pihak pemerintah daerah untuk mengatasi berbagai masalah tersebut. Upaya tersebut harus meliputi peningkatan aksesibilitas, perbaikan infrastruktur pendidikan, serta penyediaan pelatihan dan dukungan bagi tenaga pengajar agar dapat memberikan pendidikan yang berkualitas. Pentingnya pengambilan keputusan untuk mengambil tindakan segera atau bertahap dalam konteks penyediaan pendidikan berkualitas yang merata secara spasial melalui penentuan prioritas yang terinformasi (Sajjad et al., 2022). Untuk wilayah Indonesia bagian tengah dan timur membutuhkan strategi intervensi mungkin meliputi peningkatan infrastruktur pendidikan, penyediaan sumber daya tambahan, dan program-program dukungan khusus untuk siswa dan guru.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian metode k-medoids yang digunakan dalam penelitian ini berhasil mengelompokkan provinsi berdasarkan tingkat ketuntasan pendidikan yang teridentifikasi terdapat

disparitas signifikan, khususnya pada wilayah Indonesia bagian timur yaitu provinsi Papua di tingkat SD, SMP dan SMA dan pada di wilayah Indonesia tengah yaitu Provinsi di wilayah Indonesia Tengah yaitu Nusa Tenggara Timur (NTT) ditingkat SD dan SMA. Wilayah Indonesia Barat, seperti Jawa, Sumatra, dan Bali, menunjukkan tingkat ketuntasan yang lebih tinggi. Hal ini mencerminkan kesenjangan antara wilayah maju dan tertinggal dalam kualitas pendidikan. Disarankan untuk peneliti selanjutnya dapat mempertimbangkan faktor sosial ekonomi dan aksesibilitas geografis dalam analisis selanjutnya.

DAFTAR PUSTAKA

- CHANDRA, M. D., IRAWAN, E., SARAGIH, I. S., WINDARTO, A. P., & SUHENDRO, D. 2021. Penerapan Algoritma K-Means dalam Mengelompokkan Balita yang Mengalami Gizi Buruk Menurut Provinsi. *BIOS: Jurnal Teknologi Informasi Dan Rekayasa Komputer*, 2(1), 30–38. <https://doi.org/10.37148/bios.v2i1.19>
- GITA APRILIANU, E. L. H. 2022. Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Analisa Penjualan Toko Myam Hijab Penajam. *Jupiter, Vol 14 No.*
- HARDIYANTI, F., TAMBUNAN, H. S., & SARAGIH, I. S. 2019. Penerapan Metode K-Medoids Clustering Pada Penanganan Kasus Diare Di Indonesia. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1), 598–603. <https://doi.org/10.30865/komik.v3i1.1666>
- HIDAYAT, M. K., & RINA, F. 2022. Implementasi K-Means Dan K-Medoids Dalam Pengelompokan Wilayah Potensial Produksi Daging Ayam. *Jurnal Teknologi Industri Pertanian*, 32(158), 239–247. <https://doi.org/10.24961/j.tek.ind.pert.2022.32.3.239>
- JUNAEDI, E., SIREGAR, A. M., & NURLAELASARI, E. 2022. Implementasi C4.5 Dan Algoritma K-Nearest Neighbor Untuk Prediksi Kelayakan Pemberian Kredit Menggunakan RapidMiner Studio. *Scientific Student Journal for Information, Technology and Science*, 3(1), 83–90.
- KOGOYA, A., WAANI, F. J., & PAAT, C. J. 2023. Dampak Pendidikan terhadap Kualitas anak-anak Pedalaman di Kampung Mundidok Distrik Gome Utara Kabupaten Puncak Provinsi Papua. *Ilmiah Society*, 3(3), 1–6.
- M. PUJA ALIF BUDIMAN, & WINARSO, D. 2024. Penerapan Algoritma K-Medoids Clustering untuk Pengelompokan Bulan Rawan Bencana Kabut Asap di Kota Pekanbaru. *Jurnal Fasilkom*, 14(1), 1–8.
- <https://doi.org/10.37859/jf.v14i1.6858>
- MAULANA, R. 2022. Kesenjangan Mutu Pendidikan di Wilayah Timur Indonesia. *Asian Education and Development Studi*, 2(2), 1–6. https://www.researchgate.net/publication/365893465_Kesenjangan_Mutu_Pendidikan_di_Wilayah_Timur_Indonesia
- METARUM, M. F. H. 2021. Tantangan SPM: Menilik Mutu Pendidikan Sekolah Pedalaman di Ules Nusa Tenggara Timur. *Edukatif: Jurnal Ilmu Pendidikan*, 3(3), 980–988. <https://edukatif.org/index.php/edukatif/article/view/483>
- NISA, N. S., & SAMPUTRA, P. L. 2020. Analisis Ketimpangan Pendidikan Di Provinsi Papua Barat. *Jurnal Perspektif Ekonomi Darussalam*, 6(2), 115–135. <https://doi.org/10.24815/jped.v6i2.16388>
- RAHAYU, K., NOVIANTI, L., & KUSNANDAR, M. 2020. Implementation Data Mining with K-Means Algorithm for Clustering Distribution Rabies Case Area in Palembang City. *Journal of Physics: Conference Series*, 1500(1). <https://doi.org/10.1088/1742-6596/1500/1/012121>
- RICHA PUTRI, M., GIBRAN SATYA NUGRAHA, & RAMADITIA DWIYANSAPUTRA. 2023. Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Pendidikan Menggunakan Metode K-Means Clustering. *Journal of Computer Science and Informatics Engineering (J-Cosine)*, 7(1), 1–8. <https://doi.org/10.29303/jcosine.v7i1.509>
- SAJJAD, M., MUNIR, H., KANWAL, S., & NAQVI, S. A. A. 2022. Spatial inequalities in education status and its determinants in Pakistan: A district-level modelling in the context of sustainable development Goal-4. *Applied Geography*, 140.
- TAKYI, S. A., AMPONSAH, O., ASIBEY, M. O., & AYAMBIRE, R. A. 2021. An overview of Ghana's educational system and its implication for educational equity. *International Journal of Leadership in Education*, 24(2), 157–182. <https://doi.org/10.1080/13603124.2019.1613565>
- ZAHROTUL KAMALIA, A., RAHENDRA HERLIANTO, H., ROZIKIN, Z., INFORMATIKA, T., TEKNIK, F., & PELITA BANGSA, U. 2023. SIGMA-Jurnal Teknologi Pelita Bangsa Klasifikasi Rating Sekolah Menengah Atas Berdasarkan Jumlah Siswa Sekolah Di Provinsi Dki Jakarta Menggunakan Metode Algoritma Naive Bayes. *Jurnal Teknologi Pelita Bangsa*, 14(3), 167–171. <https://dapo.kemdikbud.go.id/>