

PENERAPAN ALGORITMA *K-MEANS* DAN *K-MEDOIDS* UNTUK KLASTERISASI WILAYAH BERDASARKAN INDIKATOR KUALITAS PENDIDIKAN DI PROVINSI NUSA TENGGARA TIMUR TAHUN 2023

Regita Karine Mashitawai^{1*}, Widhi Nugrahaeni Banov Putri², Fajar Dwi Cahyoko³, Muhammad Riefky⁴

^{1,2,3,4} Program Studi D-IV Demografi dan Pencatatan Sipil, Universitas Sebelas Maret, Kota Surakarta, Provinsi Jawa Tengah, Indonesia

*Penulis korespondensi: regitakarine14@student.uns.ac.id

ABSTRAK

Pendidikan merupakan fondasi utama dalam pembangunan sumber daya manusia yang berkualitas. Namun, kualitas pendidikan di Provinsi Nusa Tenggara Timur menunjukkan disparitas yang signifikan antar kabupaten/kota. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Nusa Tenggara Timur Tahun 2023 berdasarkan indikator kualitas pendidikan menggunakan algoritma *unsupervised learning*. Variabel yang digunakan dalam penelitian ini meliputi Angka Partisipasi Sekolah, Angka Partisipasi Kasar, Rata-rata Lama Sekolah, Rasio Guru-Murid SMA, Angka Melek Huruf dan Jumlah SMA yang didapatkan dari Badan Pusat Statistik (BPS) Indonesia Tahun 2023. Hasil penelitian menunjukkan bahwa dengan menggunakan algoritma *K-Means* sebagai metode yang optimum dengan *Silhouette Index* sebesar 0,459 dibandingkan dengan *K-Medoids* diperoleh hasil pengelompokan kabupaten/kota di Provinsi Nusa Tenggara Timur ke dalam dua kluster utama, yaitu kluster satu dengan kualitas pendidikan tinggi yang terdiri atas 8 kabupaten dan kluster dua dengan kualitas pendidikan rendah yang mencakup 14 kabupaten dan kota. Hasil penelitian ini dapat digunakan oleh Pemerintah Provinsi Nusa Tenggara Timur sebagai dasar penetapan prioritas kebijakan, khususnya kebijakan yang terfokus pada kabupaten/kota dengan kualitas rendah sehingga mempercepat pemerataan kualitas pendidikan di Provinsi Nusa Tenggara Timur.

Kata kunci: *K-Means*, Kualitas Pendidikan, Nusa Tenggara Timur, *Silhouette Score*.

1 PENDAHULUAN

Pendidikan merupakan fondasi utama dalam pembangunan bangsa karena berperan dalam menciptakan sumber daya manusia yang kompeten. Di Indonesia, meskipun terdapat kemajuan pada akses dan pemerataan pendidikan secara umum, faktanya kualitas serta hasil pendidikan belum merata di seluruh daerah. Ketimpangan antara daerah kota dan desa, maupun antara Indonesia bagian barat dan timur, masih menjadi kendala signifikan yang menghambat pencapaian tujuan pendidikan nasional secara menyeluruh. Provinsi Nusa Tenggara Timur (NTT) merupakan contoh wilayah yang dihadapkan dengan tantangan besar dalam pemerataan pendidikan. Kualitas pendidikan di provinsi ini berada pada peringkat rendah dibandingkan provinsi lain di Indonesia, bahkan menempati posisi keempat dari bawah (Banunu, 2021). Kondisi geografis yang terdiri dari pulau-pulau terpencil dan infrastruktur yang belum sepenuhnya memadai menjadikan akses pendidikan yang memadai masih menjadi tantangan besar. Hal ini menunjukkan bahwa akses

pendidikan di daerah terpencil Nusa Tenggara Timur tetap terbatas, dan partisipasi anak-anak pada sistem pendidikan formal belum mencapai tingkat yang optimal.

Capaian pendidikan di Nusa Tenggara Timur berkembang cukup dinamis, meskipun masih dihadapkan dengan berbagai tantangan struktural dan geografis. Beberapa indikator utama yang dapat mencerminkan capaian pendidikan di wilayah ini antara lain angka partisipasi sekolah (APS), di mana indikator ini menunjukkan persentase penduduk usia sekolah yang sedang menempuh pendidikan pada jenjang tertentu, dimana semakin tinggi jenjang pendidikan, semakin sedikit anak yang melanjutkan studi. Beberapa penyebab meliputi tingkat kemiskinan, jumlah penduduk yang menikah usia dini, dan keterbatasan akses ke pendidikan lanjutan, terutama di wilayah terpencil. Indikator angka partisipasi kasar (APK) menghitung jumlah siswa terdaftar di tingkat pendidikan tertentu tanpa mempertimbangkan batas usia yang ideal, sehingga dapat menggambarkan keberadaan siswa yang terlambat memasuki sekolah atau yang pernah berhenti dan kemudian melanjutkan kembali. APK yang tinggi namun tidak diimbangi dengan APS yang kuat menunjukkan terdapat beberapa ketimpangan dalam kesinambungan pendidikan (Yolanda & Eriani, 2022). Indikator rata-rata lama sekolah (RLS) ialah indikator yang mencerminkan jumlah tahun pendidikan formal yang telah diselesaikan oleh penduduk berusia 25 tahun ke atas. Nilai RLS di Nusa Tenggara Timur, masih rendah dibandingkan provinsi lain di Indonesia (Manurung & Hutabarat, 2021). Minimnya RLS di NTT menggambarkan bahwa banyak penduduk tidak menyelesaikan pendidikan menengah, bahkan sebagian besar hanya menamatkan pendidikan dasar. Indikator rasio jumlah guru dan murid sekolah menengah atas (SMA) menunjukkan adanya kecukupan tenaga pengajar terhadap jumlah siswa dan digunakan sebagai salah satu parameter penting dalam mengukur kualitas pendidikan. Angka melek huruf (AMH) menggambarkan persentase penduduk berusia diatas 15 tahun yang mempunyai keterampilan membaca dan menulis kalimat sederhana menggunakan huruf Latin atau aksara lainnya. Indikator jumlah SMA merupakan jumlah SMA yang tersedia di NTT menjadi tantangan besar, dikarenakan banyak desa di daerah terpencil yang belum memiliki fasilitas SMA, sehingga siswa terpaksa menempuh jarak lebih dari 10 kilometer untuk bersekolah. Hal ini berdampak pada tingginya angka siswa yang putus sekolah pada jenjang SMA.

Untuk mengelompokkan kabupaten/kota di Provinsi NTT berdasarkan indikator kualitas pendidikan seperti APS, APK, RLS, AMH, rasio jumlah guru-murid SMA, dan jumlah SMA. Pendekatan klasterisasi menggunakan algoritma *K-Means* dan *K-Medoids* ditetapkan sebagai pilihan karena keduanya mampu mengelompokkan data multivariat secara efisien dan interpretatif. Metode *K-Means* merupakan salah satu algoritma *unsupervised learning* yang sering digunakan dalam bidang pendidikan karena kemampuannya mengelompokkan data numerik dalam jumlah besar secara efisien (Basalamah & Resad, 2023). Namun, metode tersebut sensitif terhadap *outlier*, sehingga untuk memperoleh hasil klaster yang lebih stabil dan representatif, digunakan juga metode *K-Medoids* yang memiliki ketahanan lebih baik terhadap nilai ekstrem karena memanfaatkan medoid, yakni data aktual, sebagai pusat klaster. Kombinasi kedua metode tersebut mencerminkan penerapan terbaik untuk studi kasus pendidikan dengan membandingkan *K-Means* dan *K-Medoids* untuk mengelompokkan kualitas pendidikan di Indonesia yang menunjukkan bahwa *K-Medoids* memiliki keunggulan dalam stabilitas hasil meskipun membutuhkan proses komputasi lebih kompleks (Aulia & Nurahman, 2023). Melalui pendekatan kedua metode tersebut untuk wilayah di Nusa Tenggara Timur dapat dikelompokkan ke dalam klaster yang mencerminkan kesamaan kondisi pendidikan, sehingga mempermudah penyusunan kebijakan berbasis data yang lebih akurat dan efektif untuk menangani ketimpangan antar wilayah dalam pembangunan pendidikan.

Berbagai penelitian terdahulu telah memanfaatkan indikator kualitas pendidikan untuk analisis klasterisasi. Penelitian oleh Lenama et al. (2023) mengelompokkan kabupaten/kota di NTT menggunakan indikator dalam bidang pendidikan seperti APS dan angka partisipasi murni yang disimpulkan bahwa terdapat empat kelompok yang terbentuk dari penelitian ini. Kelompok satu terdiri dari 5 anggota, kelompok dua terdiri dari 7 anggota, kelompok tiga terdiri dari 8 anggota, dan kelompok empat terdiri dari 2 anggota. Penelitian lain oleh Izzuddin dan Wijayanto (2024) menggunakan indikator pendidikan seperti AMH dan RLS yang disimpulkan bahwa kualitas pendidikan di Provinsi Papua masih tergolong rendah. Selain itu penelitian oleh Khabib dan Sylvi (2024) menggunakan indikator kualitas pendidikan, APS dan RLS untuk mengelompokkan provinsi di Indonesia yang fokus pada disparitas wilayah dalam hal akses dan hasil pendidikan dapat disimpulkan bahwa terdapat empat kelompok. Kelompok satu terdiri dari 11 anggota, kelompok dua terdiri dari 17 anggota, kelompok tiga terdiri dari 6 anggota, dan kelompok empat terdiri dari 1 anggota. Penelitian terdahulu yang berkaitan dengan pendekatan *K-Means* telah digunakan oleh Simarmata et al. (2024) dalam mengelompokkan provinsi berdasarkan indeks pembangunan manusia, serta penelitian oleh Wuni et al. (2020) untuk mengelompokkan provinsi berdasarkan tingkat putus sekolah jenjang sekolah menengah pertama (SMP). Sementara itu, pendekatan *K-Medoids* telah diaplikasikan oleh Fialine et al. (2021) untuk mengelompokkan berdasarkan indikator pendidikan, serta penelitian oleh Rahmawati dan Fauzan (2024) dalam pengelompokan provinsi berdasarkan indikator pendidikan. Sementara Aulia dan Nurahman (2023) membandingkan efektivitas *K-Medoids* dan *K-Means* dalam klasterisasi data pendidikan di Kalimantan Tengah tahun 2023 dengan pendekatan optimalisasi yaitu indeks *Davies-Bouldin*. Berdasarkan uraian penelitian terdahulu menjelaskan bahwa umumnya berfokus pada *K-Medoids* atau *K-Means* secara terpisah, tanpa adanya analisis perbandingan menggunakan keenam indikator pendidikan. Penelitian ini menyandingkan algoritma *K-Means* dan *K-Medoids* secara komprehensif guna menilai kualitas pendidikan di Provinsi NTT melalui pendekatan berbasis data, sehingga mengisi celah penting dalam penilaian pendidikan. Peneliti menggunakan keenam variabel sebagai indikator kualitas pendidikan di Provinsi NTT pada tahun 2023 dengan mengelompokkan wilayah menggunakan pendekatan *K-Means* dan *K-Medoids*. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi NTT berdasarkan indikator kualitas pendidikan tahun 2023 dengan algoritma *K-Means* dan *K-Medoids*. Diharapkan hasil penelitian ini mampu memperdalam pemahaman yang lebih terstruktur dan didasarkan pada data mengenai keadaan pendidikan di Provinsi NTT, serta berperan sebagai pertimbangan bagi perumus kebijakan dalam menyusun intervensi yang lebih akurat untuk meningkatkan kualitas pendidikan di NTT.

2 METODE

Penelitian ini dilakukan dengan pendekatan kuantitatif, dimana penelitian kuantitatif merupakan suatu pendekatan terstruktur dalam mengumpulkan serta menganalisis data berbentuk angka yang bertujuan untuk menjelaskan suatu fenomena, menguji dugaan, atau meramalkan hasil tertentu (Qudratuddarsi et al., 2024). Penelitian ini memanfaatkan data sekunder, dengan sumber utama berasal dari laman resmi Badan Pusat Statistik (BPS) Provinsi NTT (<https://ntt.bps.go.id>) pada tahun 2023. Penelitian ini didasarkan pada variabel yang berkaitan dengan indikator kualitas pendidikan dilihat pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Keterangan	Satuan
X ₁	RLS	Tahun
X ₂	APS	Persen
X ₃	APK	Persen
X ₄	Rasio Guru-Murid SMA	-
X ₅	AMH	Persen
X ₆	Jumlah SMA	Persen

2.1 Tahap Penelitian

Tahapan penelitian ini dibagi menjadi enam tahapan sebagai berikut.

1. Mengumpulkan data, dimana peneliti memperoleh *dataset* dari BPS Provinsi NTT tahun 2023.
2. Mengkarakteristik data yang terkumpul menggunakan statistika deskriptif yang memuat nilai *mean*, *minimum*, *maximum*, *standard deviation*.
3. Menstandarisasi data dengan satuan yang berbeda.
4. Menentukan optimalisasi pada jumlah kelompok dengan *K-Means* dan *K-Medoids*.
5. Mengklusterkan kabupaten/kota di Provinsi NTT tahun 2023 berdasarkan indikator kualitas pendidikan dengan menggunakan metode kluster yang optimum.
6. Memberikan kesimpulan dan saran.

2.2 K-Means

K-Means merupakan algoritma *unsupervised learning* yang dikenal karena kesederhanaannya dan keefektifannya dalam pengelompokan (Robbani & Fatah, 2024). Algoritma ini bergantung pada sebuah fungsi untuk menentukan data yang memiliki karakteristik serupa. Berikut adalah langkah-langkah untuk mengelompokkan data dengan *K-Means* sebagai berikut.

1. Menentukan jumlah kluster.
2. Mengelompokkan data kedalam kluster dilakukan secara *random*.
3. Menghitung pusat kluster (*centroid*) dari data yang ada di setiap kluster (Putri et al., 2023). Jarak *Euclidean* ditulis sebagai Persamaan (1) sebagai berikut.

$$D(p, q) = \sqrt{(X_{1p} - X_{1q})^2 + (X_{2p} - X_{2q})^2 + \dots + (X_{np} - X_{nq})^2} \quad (1)$$

Keterangan :

- $D_{(p,q)}$ = Jarak data ke p ke pusat kluster q
- X_{np} = Data ke p atribut data ke n
- X_{nq} = Titik Pusat ke q pada atribut ke n

4. Mendistribusikan setiap data ke *centroid*/pusat kluster terdekat.
5. Mengulangi langkah-langkah sebelumnya sampai hasil dari pengelompokan tidak mengalami perubahan.

2.3 K-Medoids

K-Medoids merupakan algoritma pengelompokan yang dirancang khusus untuk mengidentifikasi *medoid* sebagai pusat kluster dalam proses pengelompokan data. Algoritma ini juga dikenal dengan sebutan *Partitioning Around Medoids* (PAM), yang mengacu pada pendekatannya dalam membagi data berdasarkan titik *medoid* (Luo, 2022). Dibandingkan dengan *K-Means*, *K-Medoids* memiliki keunggulan dalam hal kemudahan implementasi serta

ketahanannya terhadap *outliers*, menjadikannya pilihan yang lebih stabil dalam berbagai kasus klasterisasi.

Proses pengelompokan data pada *K-Medoids* dilakukan melalui sistem partisi, di mana setiap data *point* akan dikelompokkan berdasarkan perhitungan jarak terhadap titik *medoid*. Secara teknis, algoritma ini bekerja dengan membandingkan jarak antara data *non-medoid* (titik data biasa) dan data *medoid* (titik pusat klaster). Untuk menentukan klaster terdekat bagi suatu data, *K-Medoids* menggunakan perhitungan Jarak *Euclidean*, yang merupakan metode pengukuran jarak paling umum dalam analisis klasterisasi (Putri et al., 2023).

2.4 Silhouette Score

Silhouette Score (atau *Silhouette Coefficient*) adalah metrik evaluasi dalam analisis klastering yang digunakan untuk mengukur seberapa baik suatu objek ditempatkan dalam klasternya dibandingkan dengan klaster lain, dengan mempertimbangkan kepadatan dalam klaster dan jarak antar klaster (Pramesty et al., 2023). Berikut Persamaan (2) mengenai perhitungan *Silhouette Score*:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2)$$

Keterangan:

- $s(i)$ = *Silhouette Score* atau *Silhouette Coefficient*
- $a(i)$ = Jarak rata-rata objek ke- i ke semua objek lain dalam klaster yang sama
- $b(i)$ = Jarak rata-rata objek ke- i ke semua objek di klaster terdekat lainnya
- $\max(a(i), b(i))$ = Nilai maksimum antara $a(i)$ dan $b(i)$ untuk normalisasi

3 HASIL DAN PEMBAHASAN

3.1 Deskripsi Indikator Kualitas Pendidikan

Statistik deskriptif merupakan teknik analisis data yang berfungsi untuk memberikan gambaran umum mengenai distribusi dan karakteristik dari masing-masing indikator atau variabel dalam penelitian. Sehingga dapat dipastikan bahwa *dataset* yang digunakan telah memenuhi ketentuan dasar dan siap untuk diproses secara lebih mendalam.

Tabel 2. Descriptive Statistic

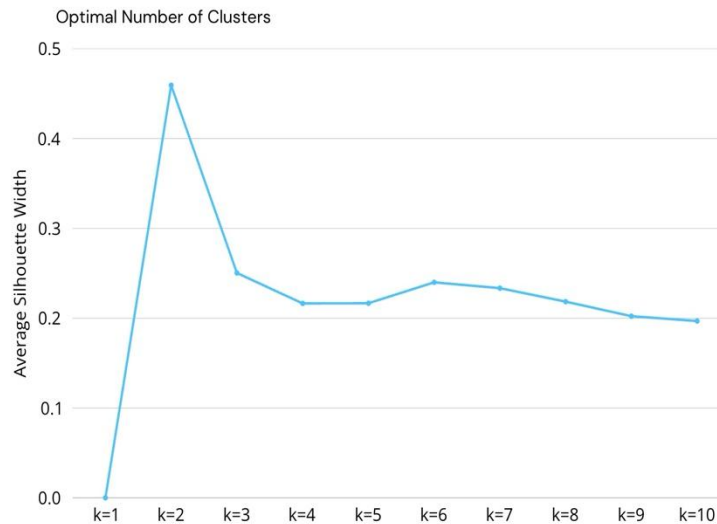
	<i>Minimum</i>	<i>Maximum</i>	<i>Mean</i>	<i>Std Deviation</i>
RLS	6.380	11.620	7.796	1.048
APS	62.910	88.590	75.080	5.914
APK	72.400	112.470	86.785	10.215
Rasio Guru-Murid SMA	7.000	17.000	12.636	2.769
AMH	90.610	99.870	97.662	2.605
Jumlah SMA	5.000	62.000	27.772	14.709

Berdasarkan Tabel 2 didapatkan karakteristik data dari setiap variabel atau indikator dalam penelitian ini. Variabel RLS di Provinsi NTT tahun 2023 memiliki rata-rata sebesar 7.796 tahun berada di dalam batas minimum sebesar 6.380 tahun dan batas maksimum sebesar 11.620 tahun dengan keragaman data sebesar 1.048 tahun. Sementara itu, variabel angka partisipasi sekolah memiliki rata-rata 75.080 persen berada di dalam batas minimum sebesar 62.910 persen dan batas maksimum sebesar 88.590 persen dengan keragaman data sebesar 5.914 persen. Variabel angka partisipasi kasar memiliki rata-rata 86.785 persen dan berada di dalam batas minimum sebesar 72.400 persen dan batas maksimum sebesar 122.470 persen dengan keragaman data sebesar 10.215

persen. Variabel rasio guru-murid SMA memiliki rata-rata sebesar 12.636 berada di dalam batas minimum sebesar 7 orang dan batas maksimum sebesar 17 orang dengan keragaman data sebesar 2.769. Adapun variabel angka melek huruf memiliki rata-rata sebesar 97.662 persen berada di dalam batas minimum sebesar 90.610 persen dan berada di dalam batas maksimum sebesar 99.870 persen dengan keragaman data sebesar 2.605 persen. Variabel jumlah sekolah memiliki rata-rata sebesar 27.772 persen dan berada di batas minimum sebesar 5 persen dan berada di batas maksimum sebesar 62 persen dengan keragaman data sebesar 14.709 persen.

3.2 K-Means

K-Means digunakan untuk klusterisasi wilayah berdasarkan indikator kualitas pendidikan dengan $k=10$. Hasil perhitungan kluster *K-Means* di *RStudio* dapat dilihat pada hasil grafik *output* yang tertera pada Gambar 1.



Gambar 1. Grafik *Silhouette Score K-Means*

Berdasarkan pada Gambar 1 terdapat analisis grafik *Silhouette Score* ditemukan jumlah kluster ideal yang terbentuk adalah dua kluster ($k=2$). Hal ini didasari oleh nilai *Silhouette Score* pada $k=2$ lebih besar dari pada nilai indeks *Silhouette* kluster lain. Berikut adalah tabel nilai *Silhouette Score* yang diperoleh:

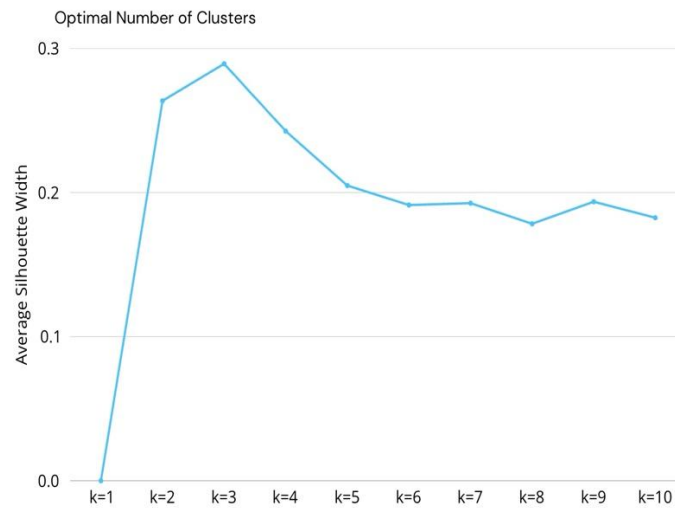
Tabel 3. *Validation Measures Algoritma K-Means*

	<i>Silhouette Score</i>
$k=2$	0.459
$k=3$	0.250
$k=4$	0.216
$k=5$	0.216
$k=6$	0.239
$k=7$	0.233
$k=8$	0.218
$k=9$	0.202
$k=10$	0.196

Berdasarkan Tabel 3 jumlah kluster paling ideal diperoleh saat $k=2$ dengan nilai *Silhouette Score* sebesar 0.459 lebih besar dari pada nilai *Silhouette Score* kluster lainnya. Nilai *Silhouette Score* memiliki rentang antara -1 hingga 1, dimana semakin besar *Silhouette Score* menunjukkan kualitas klusterisasi yang lebih baik.

3.3 K-Medoids

Proses percobaan pada tahap ini mengikuti jumlah percobaan dengan metode *K-Means*. Hasil pengelompokan yang diperoleh dari penerapan algoritma *K-Medoids* dapat diamati berdasarkan Gambar 2 sebagai berikut.



Gambar 2. Grafik *Silhouette Score K-Medoids*

Berdasarkan Gambar 2 hasil validasi menunjukkan bahwa jumlah kluster yang paling ideal pada saat $k=3$. Penentuan ini diperoleh berdasarkan dari nilai *Silhouette Score* tertinggi terletak pada kluster yang berjumlah 3.

Tabel 4. Validation Measures Algoritma *K-Medoids*

	<i>Silhouette Score</i>
$k=2$	0.263
$k=3$	0.289
$k=4$	0.242
$k=5$	0.204
$k=6$	0.191
$k=7$	0.192
$k=8$	0.178
$k=9$	0.193
$k=10$	0.182

Berdasarkan Tabel 4 menjelaskan bahwa jumlah kluster yang ideal yaitu $k=3$, ditunjukkan oleh nilai *Silhouette Score* sebesar 0.289 yang lebih besar daripada nilai *Silhouette Score* pada jumlah kluster lainnya.

3.4 Analisis Klaster Berdasarkan Metode Optimum

Penelitian ini menggunakan algoritma *K-Means* dan *K-Medoids* sebagai metode untuk mengidentifikasi jumlah klaster paling sesuai dengan karakteristik data indikator kualitas pendidikan. Penggunaan *Silhouette Score* sebagai alat verifikasi untuk membandingkan efektivitas kedua metode tersebut.

Tabel 5. Hasil Perhitungan *Silhouette Score*

	Metode	<i>Silhouette Score</i>
$k=2$	<i>K-Means</i>	0.459
$k=3$	<i>K-Medoids</i>	0.289

Berdasarkan Tabel 5 menunjukkan dengan jelas perbandingan nilai *Silhouette Score* dari kedua algoritma. Nilai *Silhouette Score* dari *K-Means* sebesar 0.459 yang terdapat pada $k=2$ lebih tinggi bila disandingkan dengan nilai *Silhouette Score* *K-Medoids*. Hal ini menunjukkan bahwa pengelompokan yang dilakukan menggunakan algoritma *K-Means* memiliki efektivitas dan akurasi yang lebih baik dibandingkan dengan *K-Medoids* dalam klasterisasi wilayah berdasarkan indikator kualitas pendidikan di Nusa Tenggara Timur. Plot hasil klaster seluruh wilayah di Provinsi NTT berdasarkan indikator kualitas pendidikan tahun 2023 diperoleh melalui penerapan algoritma *K-Means* sebagai metode yang optimum.



Gambar 3. Plot Hasil Klasterisasi Algoritma *K-Means*

Berdasarkan Gambar 3, klaster 1 ditandai dengan garis berwarna merah dan observasi berbentuk lingkaran. Sementara itu, klaster 2 ditandai dengan garis berwarna biru dan observasi berbentuk segitiga. Anggota klaster dikelompokkan berdasarkan kesamaannya dengan data lain. Hasil klasterisasi menggunakan algoritma *K-Means* berdasarkan data indikator kualitas pendidikan di Provinsi Nusa Tenggara Timur tahun 2023 berjumlah 2 klaster dan Tabel 6 menunjukkan hasil klasterisasi terbaik dengan metode yang optimum.

Tabel 6. Hasil Klasterisasi Terbaik

Klaster	Jumlah Anggota	Kabupaten/Kota
1	8	Sumba Barat, Sumba Timur, Sikka, Manggarai, Rote Ndao, Sumba Tengah, Sumba Barat Daya, dan Sabu Raijua.
2	14	Kupang, Timor Tengah Selatan, Timor Tengah Utara, Belu, Alor, Lembata, Flores Timur, Ende, Ngada, Manggarai Barat, Nagekeo, Manggarai Timur, Malaka, dan Kota Kupang

Dari data yang disajikan dalam tabel 6, dapat dilihat beberapa informasi berikut:

- a. **Klaster 1**
Pada klaster pertama terdiri dari 8 kabupaten, yaitu: Sumba Barat, Sumba Timur, Sikka, Manggarai, Rote Ndao, Sumba Tengah, Sumba Barat Daya, dan Sabu Raijua. Dari hasil pengelompokan menggunakan enam variabel tersebut menunjukkan bahwa provinsi dalam klaster satu memiliki capaian yang tergolong tinggi dalam indikator kualitas pendidikan. Capaian tersebut dipengaruhi oleh akses pendidikan yang lebih baik, penyebaran tenaga pengajar yang lebih merata dan angka partisipasi sekolah yang cukup tinggi.
- b. **Klaster 2**
Pada klaster kedua terdiri dari 14 kabupaten dan kota, yaitu: Kupang, Timor Tengah Selatan, Timor Tengah Utara, Belu, Alor, Lembata, Flores Timur, Ende, Ngada, Manggarai Barat, Nagekeo, Manggarai Timur, Malaka, dan Kota Kupang. Dari hasil pengelompokan menggunakan enam variabel tersebut menunjukkan bahwa provinsi dalam klaster satu memiliki capaian yang tergolong rendah dalam indikator kualitas pendidikan. Rendahnya capaian indikator kualitas pendidikan disebabkan oleh keterbatasan sarana dan prasarana pendidikan, kekurangan tenaga pengajar yang berkualitas serta angka partisipasi sekolah yang masih rendah terutama di wilayah terpencil.

4 KESIMPULAN

Berdasarkan penelitian yang telah dilakukan terkait data indikator kualitas pendidikan dapat disimpulkan bahwa menggunakan algoritma *K-Means* sebagai metode yang optimum dengan nilai *Silhouette Score* lebih besar yaitu 0.459 dibandingkan dengan *K-Medoids* diperoleh hasil pengelompokan kabupaten/kota di Provinsi Nusa Tenggara Timur ke dalam dua klaster utama. Klaster satu dengan kualitas pendidikan tinggi yang terdiri atas 8 kabupaten dan klaster dua dengan kualitas pendidikan rendah yang mencakup 14 kabupaten dan kota. Berdasarkan hasil penelitian yang telah dijelaskan, maka diperoleh saran untuk Pemerintah Provinsi Nusa Tenggara Timur sebagai dasar penetapan prioritas kebijakan, khususnya kebijakan yang terfokus pada kabupaten/kota dengan kualitas rendah sehingga mempercepat pemerataan kualitas pendidikan di Provinsi Nusa Tenggara Timur.

DAFTAR PUSTAKA

- Aulia, D. D., & Nurahman. (2023). Comparison Performance Of K-Medoids and K-Means Algorithms In Clustering Community Education Levels. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 12(2), 273–282.
- Banunu, N. (2021). Ketimpangan Pendidikan di Nusa Tenggara Timur: Evaluasi Kurva Kuznet dan Determinannya. *Jurnal Statiska Terapan*, 1(1), 1–20.
- Basalamah, A. T., & Resad, S. (2023). Penerapan Algoritma K-Means Clustering Pada Tingkat Penyelesaian Pendidikan Di Provinsi Indonesia. *Jurnal Informatika Dan Teknologi Komputer*, 4(2), 114–121.

- Fialine, A. P., Alodia, D. A., Endriani, D., & Widodo, E. (2021). Implementasi Metode K-Medoids Clustering untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Pendidikan. *SEPREN: Journal of Mathematics Education and Applied*, 2(2), 1–13.
- Khabib, W. 'Izza N., & Sylvi, P. (2024). Analisis Faktor Terhadap Indeks Pembangunan Manusia Dan Pengelompokan Provinsi Di Indonesia Tahun 2023. *Jurnal Saintek*, 1(1), 34–47.
- Lenama, N., Kleden, M. A., & Pasangka, I. G. (2023). K-Means Clustering Analysis Pada Pengelompokan Kabupaten/Kota Di Provinsi Nusa Tenggara Timur Berdasarkan Indiator Pendidikan. *Jurnal Cakrawala Ilmiah*, 2(9), 3365.
- Luo, X. (2022). Digital Art Design Effectiveness Model System Based on K-Medoids Algorithm. *Advances in Multimedia*, 2022, 2901815.
- Manurung, E. N., & Hutabarat, F. (2021). Pengaruh Angka Harapan Lama Sekolah, Rata-Rata Lama Sekolah, Pengeluaran per Kapita Terhadap Indeks Pembangunan Manusia. *Jurnal Ilmiah Akuntansi Manajemen*, 4(2), 121–129.
- Pramesty, M. H. E., Aprilia, Y. U., Purba, R. B. J., & Athoillah, M. (2023). Optimalisasi Kualitas Pendidikan Provinsi Nusa Tenggara Timur: Klasterisasi Kabupaten Dan Kota Menggunakan Algoritma K-Means Dengan Metode Elbow dan Shilluette Score. *Seminar Nasional Hasil Riset Dan Pengabdian*, 359–368.
- Putri, M. R., Nugraha, G. S., & Dwiyanaputra, R. (2023). Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Pendidikan Menggunakan Metode K-Means. *Journal Of Computer Science and Informatics Engineering*, 7(1), 76–83.
- Quadratuddarsi, H., Meivawati, E., Fitrasari, & Saputra, R. (2024). Pelatihan Penelitian Metode Kuantitatif dan Systematic literature review bagi Dosen dan Mahasiswa. *Jurnal Pengabdian Kepada Masyarakat*, 3(1), 22–32.
- Rahmawati, O., & Fauzan, A. (2024). Provincial Clustering Based O Education Indicators: K-Medoids Application And K-Medoids Outlier Handling. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 18(2), 1167–1178.
- Robbani, K. R., & Fatah, Z. (2024). Klasterisasi Pendidikan SD Untuk Mengetahui Daerah Dengan Pendidikan Terendah Menggunakan Algoritma K-Means. *Jurnal Ilmiah Multidisiplin Ilmu*, 1(5), 74–80.
- Simarmata, J. E., Chrisinta, D., & Purnomo, M. (2024). Implementation of K-Means Clustering to Human Development Indicators in East Nusa Tenggara. *Journal of Research in Mathematics Trends and Technology*, 6(2), 46–56.
- Wuni, S., Siregar, A. M., & Kusumaningrum, D. S. (2020). K-Means Clustering untuk Mengelompokan Tingkat Putus Sekolah Jenjang SMP di Indonesia. *Scientific Student Journal for Information, Technology and Science*, 1(1), 47–54.
- Yolanda, A. M., & Eriani, E. (2022). Optimalisasi Angka Partisipasi Sekolah Anak Usia Dini Untuk Meningkatkan Indeks Pembangunan Manusia (Studi Kasus di Provinsi Nusa Tenggara Timur). *Jurnal Statistika Terapan*, 2(1), 20–35.