

Klasterisasi Kualitas Pendidikan Dasar Berdasarkan Elemen Ekosistem Pendidikan Menggunakan K-Means DBI di Kabupaten Magelang

Ninik Tri Hartanti¹, Dina Maulina², Erni Seniwati³

^{1,2,3} Universitas Amikom Yogyakarta

Corresponding Author: Ninik Tri Hartanti, E-Mail: ninik.t@amikom.ac.id

Published: July, 2025

ABSTRAK

Salah satu pilar dasar sistem pendidikan di Indonesia adalah pendidikan dasar yang dirancang untuk membentuk kualitas sumber daya manusia di masa depan. Kategori pendidikan dasar di Indonesia meliputi jenjang Sekolah Dasar (SD) dan yang setara, serta Sekolah Menengah Pertama (SMP) dan jenjang yang setara. Kabupaten Magelang merupakan wilayah di Jawa Tengah yang memiliki perbedaan dalam ekosistem pendidikan dapat menunjukkan adanya kesenjangan dalam penyelenggaraan pendidikan dasar. Penelitian ini bertujuan untuk memetakan kondisi atau kualitas pendidikan dasar di Kabupaten Magelang berdasarkan elemen ekosistem pendidikan yaitu sekolah, guru, dan murid. Metodologi yang digunakan terdiri dari algoritma K-Means, pendekatan Elbow, dan pengukuran menggunakan Davies-Bouldin Index (DBI). Klasterisasi dilakukan dengan algoritma K-Means, jumlah klaster terbaik ditentukan melalui metode Elbow, dan keakuratan klaster divalidasi menggunakan Davies-Bouldin Index (DBI). Data yang digunakan merupakan data publik yang dapat diakses melalui <https://jateng.bps.go.id/id>. Hasil penelitian memperlihatkan bahwa beberapa sekolah tingkat pendidikan dasar di Kabupaten Magelang dapat digolongkan ke dalam 3 klaster untuk setiap sekolah, yang terbagi ke dalam klaster tingkat tinggi, menengah, dan rendah. Pada klaster tinggi, sedang dan rendah untuk SD terdapat 3 kecamatan, 12 kecamatan dan 6 kecamatan. Sedangkan untuk MI terdapat 4 kecamatan, 9 kecamatan, dan 8 kecamatan. Kategori tinggi, sedang dan rendah untuk SMP adalah 3 kecamatan, 11 kecamatan dan 7 kecamatan, dan klaster tinggi, sedang dan rendah untuk MTs adalah 3 kecamatan, 8 kecamatan dan 10 kecamatan.

Kata Kunci: *Clustering, K-Means, Davies-Bouldin Index, Ekosistem Pendidikan*

ABSTRACT

One of the basic pillars of the education system in Indonesia is basic education which is designed to shape the quality of human resources in the future. The category of basic education in Indonesia includes Elementary School (SD) and equivalent levels, as well as Junior High School (SMP) and equivalent levels. Magelang Regency is an area in Central Java that has differences in the education ecosystem which can indicate gaps in the implementation of basic education. This study aims to map the conditions or quality of basic education in Magelang Regency based on the elements of the education ecosystem, namely schools, teachers, and students. The methodology used consists of the K-Means algorithm, the Elbow approach, and measurement using the Davies-Bouldin Index (DBI). Clustering is carried out using the K-Means algorithm, the number of best clusters is determined using the Elbow method, and the accuracy of the cluster is validated using the Davies-Bouldin Index (DBI). The data used is public data that can be accessed via <https://jateng.bps.go.id/id>. The results of the study show that several elementary schools in Magelang Regency can be classified into 3 clusters for each school, which are divided into high, medium, and low level clusters. In the high, medium and low clusters for elementary schools there are 3 sub-districts, 12 districts and 6 sub-districts. While for MI there are 4 sub-districts, 9 sub-districts, and 8 sub-districts. The high, medium and low categories for junior high schools are 3 sub-districts, 11 sub-districts and 7 sub-districts, and the high, medium and low clusters for MTs are 3 sub-districts, 8 sub-districts and 10 sub-districts.

Keywords: *Clustering, K-Means, Davies-Bouldin Index, Educational Ecosystem*

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0>

PENDAHULUAN

Jenjang SD dan SMP maupun yang setara merupakan cakupan dari pendidikan dasar di Indonesia dalam proses untuk menghasilkan kualitas sumber daya manusia yang terbaik karena peran^{AN} pendidikan dasar akan menjadi tahap awal dalam pembentukan pengetahuan, keterampilan, dan karakter setiap murid. Perkembangan pendidikan akan sejalan dengan perkembangan jaman yang didorong dengan perkembangan teknologi. Semakin maju teknologi, maka pendidikan juga harus mengalami perkembangan sehingga mampu untuk menghadapi tantangan jaman. Sebagai bagian dari wilayah di Jawa tengah, kabupaten Magelang menghadapi tantangan dalam mewujudkan pemerataan pendidikan dasar akibat perbedaan sosial, ekonomi, dan geografis yang cukup beragam, sehingga sangat diperlukan adanya pemerataan dan peningkatan dalam penyelenggaraan sistem pendidikan dasar. Hal ini dikarenakan adanya perbedaan faktor sosial, ekonomi, dan geografis yang beraneka ragam. Pendekatan yang dapat digunakan untuk mengantisipasi kondisi ini adalah dengan memperhatikan faktor ekosistem pendidikan yang terdiri dari sarana prasarana, tenaga pengajar, murid, peranan masyarakat, kebijakan dan teknologi.

Salah satu langkah dan upaya dalam persiapan untuk menghadapi tantangan jaman adalah pemenuhan kebutuhan pendidikan yang akan menciptakan ekosistem pendidikan yang sesuai. Elemen dalam ekosistem pendidikan yang diperlukan diantaranya adalah ketersediaan gedung sekolah, meliputi sarana dan prasarana yang memadai, ketersediaan tenaga pendidik atau guru yang berkompentensi, dan murid untuk setiap jenjang pendidikan (Juita et al., 2024).

Kabupaten Magelang, berdasarkan data <https://pusaka.magelangkab.go.id/data> memiliki luas lahan 1.085,73 km², mencakup 21 kecamatan dengan ketinggian wilayah yang berada antara 202 hingga 1.378 meter di atas permukaan laut. Kecamatan Kajoran merupakan proporsi wilayah kecamatan terluas mencapai 7,68% dari total area Kabupaten Magelang, sedangkan kecamatan dengan luas terkecil di Kabupaten Magelang adalah Ngluwar, yaitu 22,44 km² atau setara 2,07% dari total luas kabupaten Magelang. Penyebaran ekosistem pendidikan di kabupaten Magelang untuk setiap kecamatan masih dirasa kurang dikarenakan beberapa faktor diantaranya adalah kondisi lingkungan masyarakat dan ketersediaan dari elemen ekosistem pendidikan itu sendiri.

Octavia Rahmawati dan Achmad Fauzan menggunakan teknik K-Medoids untuk mengklaster provinsi di Indonesia berdasarkan indikator pendidikan. Proses evaluasi kualitas klaster, digunakan Silhouette Index dengan hasil bervariasi tergantung penanganan outlier, dan terdapat tiga teknik penanganan data yang dilakukan. Pada teknik tanpa penanganan outlier, terbentuk 2 klaster dengan nilai Silhouette Index sebesar 0,24. Sementara itu, penanganan outlier menggunakan metode mean imputation menghasilkan 8 klaster dengan nilai Silhouette Index sebesar 0,26, dan metode min-max imputation menghasilkan 9 klaster dengan nilai Silhouette Index sebesar 0,25. Hasil klasterisasi dapat digunakan sebagai dasar bagi pemangku kebijakan untuk merancang program peningkatan kualitas pendidikan yang lebih tepat sasaran di berbagai provinsi (Rahmawati & Fauzan, 2024).

Widya Syaharani dan Sriani melakukan penelitian yang bertujuan untuk mengelompokkan data penduduk miskin di Desa Sei Litur Tasik, yang terdiri dari 426 kepala keluarga, menggunakan algoritma K-Means dalam rangka mendukung program pengentasan kemiskinan ekstrem. Data yang dianalisis mencakup informasi mengenai profil penduduk, jenis pekerjaan kepala keluarga, kepemilikan tabungan, kepemilikan rumah, kondisi fisik dinding rumah, serta risiko stunting. Hasil evaluasi klasterisasi melalui Davies-Bouldin Index menunjukkan bahwa jumlah klaster paling optimal adalah dua, dengan nilai DBI sebesar 0,16. Hasil klasterisasi dapat digunakan sebagai bahan acuan bagi pemerintah dalam penyelenggaraan program pengentasan kemiskinan sehingga bisa efektif dan tepat sasaran (Syaharani et al., 2024).

Penelitian Irwan et al mengelompokkan wilayah yang terkena dampak banjir di Kabupaten Bone tahun 2020-2021 dengan menggunakan metode *K-Means Clustering* yang kemudian dengan menggunakan aplikasi *ArcGIS* hasil klastering divisualisasikan dalam bentuk peta spasial. Pada awal proses klastering ditentukan menjadi 3, 4 dan 5 klaster. Selanjutnya dengan menggunakan *Davies-Bouldin Index* (DBI) dihasilkan tahun 2020 nilai DBI untuk 3 klaster = 0.92 untuk 4 klaster = 0.73 dan untuk 5 klaster = 0.79, sedangkan tahun 2021 nilai DBI untuk 3 klaster = 0.75 untuk 4 klaster = 0.44 dan untuk 5 klaster = 0.71. Berdasarkan hasil tersebut, klaster terbaik adalah 4, sehingga terdapat 4 kategori untuk klasterisasinya, yaitu 34 desa klaster dampak ringan, 7 desa klaster dampak sedang, 2 desa klaster dampak parah dan 3 desa klaster dampak cukup tinggi (Irwan et al., 2023). Penelitian Rui Liu mengelompokkan siswa berdasarkan penilaian komprehensif, melakukan validasi keefektifan K-Means dengan menggunakan metode *DRCluster* dan *MiCluster*. Tahapan dalam penelitian meliputi integrasi data, seleksi data, pembersihan, dan transformasi data kemudian proses iteratif dengan algoritma K-Means. Dihasilkan akurasi K-Means mencapai 95.6%, lebih tinggi 12.1% dari *Mi Cluster* dan 6.8% dari *DRCluster*. Penggunaan model ini menunjukkan peningkatan skor siswa setelah dilakukan manajemen pengajaran yang ditargetkan berdasarkan hasil klasterisasi (Liu, 2022). Dalam penelitian Dwi Yanto et al tahun 2023 menerapkan algoritma *K-Means* untuk mengelompokkan kualitas sekolah dasar di Kabupaten Probolinggo yang terdiri dari 573 sekolah dasar negeri berdasarkan indikator internal dan eksternal. Indikator internal meliputi literasi, numerasi, dan karakter. Sedangkan indikator eksternal meliputi refleksi pembelajaran, kepemimpinan, dan berbagai iklim sekolah. Dengan bantuan software *RapidMiner* dihasilkan 3 klaster yaitu 290 sekolah untuk klaster unggul, 75 sekolah untuk klaster berkembang dan 208 sekolah untuk klaster kurang baik. Penelitian Dwi Yanto ini melakukan pendekatan berbasis data (*evidence-based*) untuk mendukung kebijakan, sehingga perlu adanya pembinaan pada sekolah dengan kualitas yang kurang baik (Yanto et al., 2023).

Penelitian yang dilakukan oleh Annisa Zuhri Apridayanti et al adalah mengklasifikasikan semua provinsi di Indonesia berdasarkan indikator pendidikan, dengan memanfaatkan metode K-Means Clustering untuk menentukan jumlah klaster yang paling optimal serta menggunakan Davies-Bouldin Index (DBI) sebagai metode validasinya. Data yang digunakan berasal dari Kemendikbudristek RI tahun 2021. Dengan menggunakan DBI dihasilkan nilai DBI = 1.31 pada klaster = 5. Pada klaster 1 terdapat 1 provinsi (DKI Jakarta) dengan nilai tertinggi untuk prosentase putus sekolah, prosentase lama sekolah, dan ruang kelas layak. Klaster 2 terdapat 15 provinsi dengan nilai sedang untuk indikator yang dominan di Jawa dan Kalimantan, Klaster 3 terdapat 9 provinsi dengan nilai tinggi untuk laboratorium dan perpustakaan. Di klaster 4 terdapat 8 provinsi dengan nilai tinggi dalam angka partisipasi kasar, sedangkan di klaster 5 hanya terdapat 1 propinsi (Papua) dengan rasio murid per guru yang besar. Hasil penelitian klasterisasi tersebut dapat digunakan sebagai dasar penyusunan kebijakan pemerintah terkait penyelenggaraan pemerataan kualitas pendidikan di tingkat nasional (Apridayanti et al., 2024).

Tujuan dari penelitian ini adalah mengelompokkan wilayah di Kabupaten Magelang berdasarkan tingkat kualitas pendidikan dasar dengan berdasarkan elemen ekosistem pendidikan yaitu sekolah, guru dan murid sebagai indikatornya, memanfaatkan *K-Means*, *Elbow* dan DBI, yang selanjutnya dapat digunakan untuk menganalisis elemen ekosistem pendidikan tersebut terkait ketersediaan tenaga pendidik atau guru, gedung sekolah, serta ketersediaan murid dalam suatu wilayah. Metode yang digunakan untuk klastering adalah algoritma *K-Means*, dengan menggunakan metode *Elbow* untuk menetapkan jumlah klaster yang optimal, serta memanfaatkan Davies-Bouldin Index (DBI) dalam mengevaluasi kualitas klaster yang dihasilkan. Merujuk pada fakta-fakta atau informasi yang diperoleh setelah melakukan penelitian, diharapkan dapat meningkatkan kesadaran *stakeholder* lokal seperti

sekolah, masyarakat, dan pemerintah daerah setempat terkait pentingnya memperkuat elemen ekosistem pendidikan secara menyeluruh, sehingga dapat membantu pihak berwenang seperti Dinas Pendidikan Kabupaten Magelang dalam mengidentifikasi wilayah berdasarkan kualitas elemen ekosistem pendidikan, serta dapat menjadi masukan bagi pemerintah daerah setempat dalam merumuskan kebijakan yang lebih tepat sasaran dalam meningkatkan mutu pendidikan dasar.

METODE PENELITIAN

Berdasarkan uraian di atas, pada tahap ini adalah merupakan langkah-langkah dari penelitian yang dilakukan. Sebagai langkah awal, penelitian diawali dengan identifikasi masalah, kemudian dilakukan proses pengumpulan data dan pra pemrosesan data. Langkah tersebut akan memberikan dan menentukan data yang akan digunakan. Selanjutnya adalah menerapkan metode Elbow untuk menentukan banyaknya kluster yang terbai, dilanjutkan ke proses perhitungan *K-Means* dengan dilakukan proses evaluasi kluster menggunakan DBI (*Davies-Bouldin Index*).

Identifikasi Masalah

Pentingnya kualitas pendidikan dan tantangan yang dihadapi maka perlu adanya keseimbangan untuk penyelenggaraan elemen dalam ekosistem pendidikan. Dunia pendidikan khususnya pendidikan dasar sangat berperan penting demi terbentuknya sumber daya manusia yang tangguh demi masa depan.

Pengumpulan Data

Proses ini melibatkan pengumpulan data dengan cara mengakses data publik melalui Dapodik, BPS dan data pendidikan <https://jateng.bps.go.id/id>. Data yang digunakan adalah data sekolah tahun 2023 2024 untuk pendidikan dasar, yaitu jenjang sekolah dasar dan menengah pertama seperti SD, MI, SMP, serta MTs yang ada di kabupaten Magelang. Adapun data elemen ekosistem pendidikan yang digunakan sebagai indikatornya adalah data sekolah, data guru, dan data murid.

Pra Pemrosesan Data

Di tahap ini dilakukan normalisasi data dengan menggunakan *Max-Min Scalling*, sehingga data akan bernilai biner. Tujuan proses normalisasi dalam *K-Means* yaitu dapat memberikan hasil yang valid. Tahap ini bertujuan untuk menghindari dominasi fitur dengan skala besar, meningkatkan akurasi dan stabilitas klasterisasi, dan membantu konvergensi lebih cepat (Tu et al., 2024).

$$X' = (X - X_{\min}) / (X_{\max} - X_{\min}) \quad \text{persamaan (1)}$$

Keterangan

X' : nilai setelah dinormalisasi

X : nilai asli yang akan dihitung

$X_{\min}$: nilai minimal yang ada dalam dataset

$X_{\max}$: nilai maksimal yang ada dalam dataset

Penentuan Jumlah Kluster dengan metode Elbow

Pada tahap ini, jumlah kluster dihasilkan dengan memanfaatkan metode Elbow sebagai acuan, dan merupakan metode grafis yang berfungsi menemukan nilai k terbaik pada proses *K-Means clustering*. Pada prosesnya akan terdapat beberapa perhitungan *K-Means* dengan nilai k yang berbeda-beda untuk jumlah klasternya, kemudian menghitung *Sum of Squared Errors* (SSE) untuk setiap nilai k, setelah itu mencari suatu titik yang membentuk seperti siku pada grafik yang dihasilkan. Jumlah kluster optimal dapat dilihat dari titik siku pada grafik tersebut (Permadi et al., 2023).

$$SSE = \sum_k^k = 1 \sum_{x \in S_k} \|x_1 - C_k\|_2^2 \quad \text{persamaan (2)}$$

Nilai k merepresentasikan kluster ke- c , sementara x_i menunjukkan jarak antara objek ke- i dan C_k yang merupakan pusat kluster ke- i . Melalui perhitungan dengan bantuan Elbow, banyaknya kluster yang terbaik dapat ditentukan untuk mendukung proses klusterisasi menggunakan algoritma K-Means (Hartanti et al., 2024)

Proses Klusterisasi dengan K-Means

Setelah melewati sejumlah tahapan sebelumnya, tahap selanjutnya adalah klusterisasi dengan menggunakan algoritma K-Means, yang termasuk dalam jenis *unsupervised learning*, yang mampu mengelola data tanpa memerlukan label kategori. Adapun keuntungan *K-Means* adalah relatif mudah diimplementasikan, konvergen dan memungkinkan penentuan awal posisi *centroid*. Sedangkan kelemahan dari *K-Means* adalah sensitif terhadap kondisi awal, dan tidak mampu untuk menangani data kategoris (Nurahman & Aulia, 2023).

Tahapan dalam perhitungan *K-Means* (Luchia et al., 2022)

- 1) Proses diawali dengan mengalokasikan setiap data ke kluster terdekat, berdasarkan perhitungan jarak antara data dan pusat kluster.
- 2) Memperbarui pusat kluster berdasarkan anggota kluster yang telah terbentuk sebelumnya.
- 3) Proses pengelompokan untuk setiap data dilakukan kembali dengan menggunakan pusat kluster yang telah diperbarui. Langkah ini diulangi hingga pusat kluster mencapai kondisi stabil atau konvergen.
- 4) Ulangi langkah-langkah tersebut secara terus-menerus sampai tidak terjadi perubahan lagi pada posisi pusat kluster atau konvergen.

Untuk menghitung jarak antara data dan pusat kluster, digunakan metode *Euclidean Distance* berikut:

$$d(P,Q)=\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad \text{persamaan (3)}$$

Keterangan

P dan Q: dua buah titik atau vektor dalam ruang berdimensi n

p_i : nilai atribut ke- i dari titik P

q_i : nilai atribut ke- i dari titik Q

$(p_i - q_i)^2$: selisih kuadrat dari masing-masing atribut

Evaluasi Kluster dengan DBI

Tahap selanjutnya adalah evaluasi kluster dengan menggunakan DBI (*Davies-Bouldin Index*) yang menghitung rasio antara jarak antar kluster dan jarak intra kluster. Jarak antar kluster mengukur seberapa jauh pusat kluster satu dengan lainnya, sedangkan jarak intra kluster mengukur seberapa dekat titik data dalam satu kluster (Oktavianto & Fitriyah, 2021). Persamaan untuk menentukan nilai DBI adalah:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \left(\frac{S_i + S_j}{M_{ij}} \right)$$

Keterangan

k : jumlah kluster

S_i : rata-rata jarak antara setiap titik dalam kluster ke- i terhadap pusat klusternya

S_j : rata-rata jarak dalam kluster ke- j

M_{ij} : jarak antara pusat kluster i dan pusat kluster j

$$\frac{S_i + S_j}{M_{ij}} : \text{rasio antara dispersi dalam kluster dan jarak antar kluster}$$

Semakin kecil nilai DBI, maka hasil klusterisasi semakin baik, yang menandakan bahwa kluster tidak menyebar terlalu jauh dan antar kluster terpisah. Sebaliknya, jika nilai DBI besar maka kluster tidak jelas pemisahannya. Setelah melalui beberapa tahapan dalam penelitian di atas, akan diperoleh hasil klustering yang jumlah klasternya telah optimal.

HASIL DAN PEMBAHASAN

Pengumpulan Data dan Pra Pemrosesan Data

Tahapan selanjutnya adalah hasil dari penelitian. Berdasarkan langkah-langkah dalam penelitian maka hasil penelitian dapat dijabarkan sesuai langkahnya. Berawal dari penentuan data yang digunakan seperti pada tabel 1 dan tabel 2.

Tabel 1. Data Ekosistem Pendidikan Dasar yang digunakan

Kecamatan	Jenis Sekolah	Jumlah Sekolah	Jumlah Guru	Jumlah Murid
Salaman	SD	35	303	4319
	MI	17	157	2258
	SMP	14	170	2580
	MTs	5	140	2319
Borobudur	SD	31	286	3384
	MI	15	129	2031
	SMP	7	113	1587
	MTs	2	67	991
Ngluwar	SD	20	178	1984
	MI	7	53	478
	SMP	3	59	868
	MTs	2	11	136
Salam	SD	22	190	2529
	MI	13	116	1314
	SMP	7	124	2162
	MTs	2	18	170
Srumbung	SD	29	264	2619
	MI	16	95	1361
	SMP	7	95	1730
	MTs	2	42	585
Dukun	SD	27	247	2435
	MI	8	52	934
	SMP	5	69	1014
	MTs	5	82	1182
Muntilan	SD	38	427	6789
	MI	9	104	1590
	SMP	13	235	3783
	MTs	5	63	628
Mungkid	SD	27	282	4135
	MI	17	122	2070
	SMP	6	235	2308
	MTs	5	60	486
Sawangan	SD	32	282	3618
	MI	10	58	1169
	SMP	8	124	1957
	MTs	3	25	172
Candimulyo	SD	26	215	2899

	MI	13	81	959
	SMP	4	90	1458
	MT's	1	16	126
Mertoyudan	SD	42	421	6249
	MI	12	131	2451
	SMP	9	187	3378
	MT's	1	31	279
Tempuran	SD	26	243	3591
	MI	11	107	1256
	SMP	8	109	1768
	MT's	3	41	405
Kajoran	SD	29	226	3292
	MI	24	165	2055
	SMP	4	82	1274
	MT's	7	82	787
Kaliangkrik	SD	24	186	3034
	MI	21	139	2392
	SMP	4	57	1016
	MT's	3	81	1006
Bandongan	SD	29	252	3577
	MI	17	129	2273
	SMP	6	125	2714
	MT's	4	49	368
Windusari	SD	21	193	2743
	MI	17	126	1856
	SMP	3	47	890
	MT's	5	89	877
Secang	SD	28	264	4178
	MI	23	189	2962
	SMP	5	113	1962
	MT's	8	125	1718
Tegalrejo	SD	27	239	3652
	MI	17	109	1608
	SMP	5	98	2402
	MT's	3	90	1171
Pakis	SD	31	264	3598
	MI	9	73	958
	SMP	7	100	1500
	MT's	1	26	501
Grabag	SD	36	340	5586
	MI	27	230	3676
	SMP	5	109	1839
	MT's	5	153	2169
Ngablak	SD	23	173	2537
	MI	10	59	997
	SMP	4	58	698
	MT's	2	58	646

Merujuk tabel 1 diketahui bahwa terdapat 21 kecamatan di kabupaten Magelang, dan terdapat 4 jenis sekolah sebagai pendidikan dasar yaitu SD, MI, SMP dan MT's baik yang status sekolah negeri maupun swasta. Adapun indikator sebagai elemen dalam ekosistem pendidikan yang digunakan adalah jumlah sekolah, jumlah guru termasuk tenaga pendidikan, serta jumlah murid untuk setiap jenis sekolah. Langkah selanjutnya adalah pra pemrosesan data melakukan normalisasi data dengan menggunakan *min-max scaling* seperti pada persamaan (1). Hasil normalisasi sekolah dengan menggunakan persamaan (1) seperti pada tabel 2.

Tabel 2. Data Normalisasi Sekolah SD MI SMP dan MT's

Kecamatan	SD	MI	SMP	MT's
Salaman	0,682	0,500	1,000	0,571
Borobudur	0,500	0,400	0,364	0,143
Ngluwar	0,000	0,000	0,000	0,143
Salam	0,091	0,300	0,364	0,143
Srumbung	0,409	0,450	0,364	0,143
Dukun	0,318	0,050	0,182	0,571
Muntilan	0,818	0,100	0,909	0,571
Mungkid	0,318	0,500	0,273	0,571
Sawangan	0,545	0,150	0,455	0,286
Candimulyo	0,273	0,300	0,091	0,000
Mertoyudan	1,000	0,250	0,545	0,000
Tempuran	0,273	0,200	0,455	0,286
Kajoran	0,409	0,850	0,091	0,857
Kaliangkrik	0,182	0,700	0,091	0,286
Bandongan	0,409	0,500	0,273	0,429
Windusari	0,045	0,500	0,000	0,571
Secang	0,364	0,800	0,182	1,000
Tegalrejo	0,318	0,500	0,182	0,286
Pakis	0,500	0,100	0,364	0,000
Grabag	0,727	1,000	0,182	0,571
Ngablak	0,136	0,150	0,091	0,143

Hasil normalisasi guru terlampir pada tabel 3, dan hasil normalisasi murid ada di tabel 4.

Tabel 3. Data Normalisasi Guru SD MI SMP dan MT's

Kecamatan	SD	MI	SMP	MT's
Salaman	0,51	0,59	0,65	0,91
Borobudur	0,44	0,43	0,35	0,39
Ngluwar	0,02	0,01	0,06	0,00
Salam	0,07	0,36	0,41	0,05
Srumbung	0,36	0,24	0,26	0,22
Dukun	0,29	0,00	0,12	0,50
Muntilan	1,00	0,29	1,00	0,37
Mungkid	0,43	0,39	0,52	0,35
Sawangan	0,43	0,03	0,41	0,10
Candimulyo	0,17	0,16	0,23	0,04
Mertoyudan	0,98	0,44	0,74	0,14
Tempuran	0,28	0,31	0,33	0,21
Kajoran	0,21	0,63	0,19	0,50
Kaliangkrik	0,05	0,49	0,05	0,49
Bandongan	0,31	0,43	0,41	0,27
Windusari	0,08	0,42	0,00	0,55
Secang	0,36	0,77	0,35	0,80
Tegalrejo	0,26	0,32	0,27	0,56
Pakis	0,36	0,12	0,28	0,11
Grabag	0,66	1,00	0,33	1,00
Ngablak	0,00	0,04	0,06	0,33

Tabel 4. Data Normalisasi Murid SD MI SMP dan MT's

Kecamatan	SD	MI	SMP	MT's
Salaman	0,486	0,557	0,610	1,000
Borobudur	0,291	0,486	0,288	0,394

Ngluwar	0,000	0,000	0,055	0,005
Salam	0,113	0,261	0,475	0,020
Srumbung	0,132	0,276	0,335	0,209
Dukun	0,094	0,143	0,102	0,482
Muntilan	1,000	0,348	1,000	0,229
Mungkid	0,448	0,498	0,522	0,164
Sawangan	0,340	0,216	0,408	0,021
Candimulyo	0,190	0,150	0,246	0,000
Mertoyudan	0,888	0,617	0,869	0,070
Tempuran	0,334	0,243	0,347	0,127
Kajoran	0,272	0,493	0,187	0,301
Kaliangkrik	0,219	0,598	0,103	0,401
Bandongan	0,332	0,561	0,653	0,110
Windusari	0,158	0,431	0,062	0,342
Secang	0,457	0,777	0,410	0,726
Tegalrejo	0,347	0,353	0,552	0,477
Pakis	0,336	0,150	0,260	0,171
Grabag	0,750	1,000	0,370	0,932
Ngablak	0,115	0,162	0,000	0,237

Data normalisasi di atas merupakan data pokok yang akan digunakan dalam tahap perhitungan selanjutnya, yaitu perhitungan dengan menggunakan algoritma K-Means untuk memperoleh hasil akhir yaitu kelompok berdasarkan kategori kluster tinggi, sedang dan rendah.

Penerapan metode Elbow

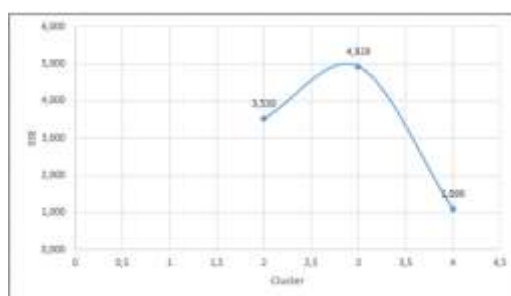
Langkah berikutnya adalah proses penentuan jumlah kluster dengan menggunakan metode *Elbow*. Pada tahap ini dilakukan sebanyak 3 kali uji coba yang setiap ujicoba dilakukan dengan uji 2 kluster, uji 3 kluster dan uji 4 kluster. Berdasarkan persamaan (2) akan dihasilkan data terlampir tabel 5.

Tabel 5. Data Nilai SSE untuk SD

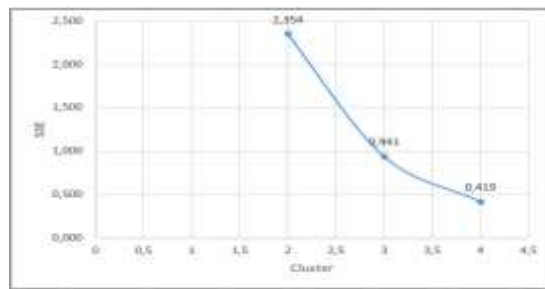
Uji Coba	Jml data	2 Kluster	3 Kluster	4 Kluster
1	21	3,530	4,929	1,099
2	15	2,354	0,941	0,419
3	17	3,465	0,974	0,704

Merujuk tabel 5 diketahui bahwa setiap hasil uji coba akan memberikan nilai SSE yang berbeda untuk setiap penentuan klusternya. Uji coba 1 dengan jumlah data 21 untuk penentuan 2 kluster nilai SSE = 3.530 dengan 3 kluster nilai SSE = 4.929 dan dengan penentuan 4 kluster nilai SSE = 1.099. Uji coba 2 dengan jumlah data 15 untuk penentuan 2 kluster nilai SSE = 2.354 dengan 3 kluster nilai SSE = 0.941 dan dengan penentuan 4 kluster nilai SSE = 0.419. Sedangkan di Uji coba 3 jumlah data 17 pada penentuan 2 kluster nilai SSE = 3.465 dengan 3 kluster nilai SSE = 0.974 dan dengan penentuan 4 kluster nilai SSE = 0.704. Berdasarkan hasil tersebut, dapat diperoleh grafik seperti gambar 1.

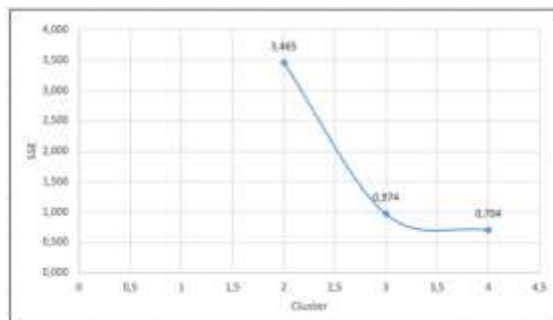
Gambar 1. Grafik Nilai SSE Uji Coba1-SD



Gambar 2. Grafik Nilai SSE Uji Coba2-SD



Gambar 3. Grafik Nilai SSE Uji Coba3-SD



Hasil nilai SSE pada gambar 1, 2 dan 3 menunjukkan bahwa lengkungan yang menyerupai siku pada grafik berada di angka 3, sesuai dengan pengertian dari metode *Elbow* maka angka yang membentuk siku tersebut merupakan jumlah kluster optimal. Langkah ini juga dilakukan pada tahap penentuan jumlah kluster untuk sekolah MI, SMP dan MTs. Dari semua data ujicoba1, 2 dan 3 baik untuk sekolah MI, SMP dan MTs diperoleh nilai $k=3$, sehingga hasil perhitungan untuk penentuan jumlah kluster adalah 3.

Penerapan Algoritma K-Means

Merujuk hasil penentuan jumlah kluster dengan metode Elbow dihasilkan nilai $k=3$, selanjutnya adalah proses perhitungan dengan algoritma K-Means untuk menentukan kelompok berdasarkan kategori per klasternya. Data untuk proses penentuan klastering dengan algoritma K-Means adalah data terlampir pada tabel 2, 3 dan 4 yaitu data normalisasi sekolah, normalisasi guru dan normalisasi murid. Dari data SD proses klastering dengan *K-Means* ditentukan nilai *centroid* adalah data wilayah Muntilan, Kaliangkrik dan Ngablak, terlampir pada Tabel 4.

Tabel 6. Centroid awal untuk SD

Centroid	Kecamatan	Jml_Sekolah	Jml_Guru	Jml_Murid
1	Muntilan	0,818	1,000	1,000
2	Kaliangkrik	0,182	0,051	0,219
3	Ngablak	0,136	0,000	0,115

Selanjutnya dengan menerapkan persamaan (3), berdasarkan data pada Tabel 6 akan diperoleh hasil terlampir pada Tabel 7.

Tabel 7. Daftar Iterasi dengan K-Means untuk SD

Iterasi	klaster1	klaster2	klaster3
1	4	13	4
2	4	12	5
3	3	13	5
4	3	12	6
5	3	12	6
6	3	12	6

Merujuk Tabel 7 diperoleh hasil bahwa terdapat 3 kluster dengan melalui 6 iterasi, data kelompok tidak berubah. Hasil iterasi menunjukkan bahwa di klaster1 terdapat 3 wilayah dengan kategori kluster tinggi dilihat dari ketersediaannya elemen ekosistem pendidikan di Sekolah Dasar (SD). Untuk kluster sedang, menunjukkan ketersediaannya elemen ekosistem pendidikan di Sekolah Dasar masih kategori sedang terdapat di 12 wilayah dan 6 wilayah untuk kategori kluster rendah yang ketersediaannya elemen ekosistem pendidikan di Sekolah Dasar masih rendah. Adapun wilayah yang masuk dalam kluster tinggi yaitu Muntilan, Mertoyudan dan Grabag. Wilayah dengan kluster sedang yaitu Salaman, Borobudur, Srumbung, Dukun, Mungkid, Sawangan, Tempuran, Kajoran, Bandongan, Secang, Tegalrejo dan Pakis. Wilayah dengan kluster rendah yaitu Ngluwar, Salam, Candimulyo, Kaliangkrik, Windusari dan Ngablak.

Tabel 8. Centroid awal untuk MI

Centroid	Kecamatan	Jml_Sekolah	Jml_Guru	Jml_Murid
1	Muntilan	0,818	1,000	1,000
2	Kaliangkrik	0,182	0,051	0,219
3	Ngablak	0,136	0,000	0,115

Tabel 9. Centroid awal untuk SMP

Centroid	Kecamatan	Jml_Sekolah	Jml_Guru	Jml_Murid
1	Muntilan	0,909	1,000	1,000
2	Kaliangkrik	0,091	0,053	0,103
3	Ngablak	0,091	0,059	0,000

Tabel 10. Centroid awal untuk MTs

Centroid	Kecamatan	Jml_Sekolah	Jml_Guru	Jml_Murid
1	Salaman	0,571	0,908	1,000
2	Mungkid	0,571	0,345	0,164
3	Ngablak	0,143	0,331	0,237

Tabel 8, 9 dan 10 merupakan data penentuan centroid awal dalam proses perhitungan dengan *K-Means*. Dalam tabel menunjukkan ada centroid yang menggunakan wilayah atau kecamatan yang sama, yaitu penentuan nilai centroid awal untuk SD, MI, dan SMP. Sedangkan wilayah yang berbeda digunakan untuk nilai penentuan centroid awal pada sekolah MTs. Hal ini dilakukan untuk memberikan nilai valid dalam penentuan nilai centroid awal di semua iterasi. Sehingga diperoleh hasil seperti pada tabel 11, 12 dan 13.

Tabel 11. Daftar Iterasi dengan K-Means untuk MI

Iterasi	klaster1	klaster2	klaster3
1	5	8	8
2	4	9	8
3	4	9	8
4	4	9	8

Tabel 12. Daftar Iterasi dengan K-Means untuk SMP

Iterasi	klaster1	klaster2	klaster3
1	3	17	1
2	3	13	5
3	3	11	7
4	3	11	7

Tabel 13. Daftar Iterasi dengan K-Means untuk MT's

Iterasi	klaster1	klaster2	klaster3
1	3	6	12
2	3	8	10
3	3	8	10
4	3	8	10

Merujuk data dari ketiga tabel di atas, dapat diperoleh hasil bahwa setiap proses perhitungan K-Means melibatkan sebanyak 4 iterasi sehingga dapat ditentukan kategorinya berdasarkan ketersediaan jumlah elemen ekosistem pendidikan dalam wilayah atau kecamatan menjadi 3 klaster. Klaster1 untuk kategori tinggi, klaster2 untuk kategori sedang, dan klaster 3 untuk kategori rendah.

KESIMPULAN

Penelitian yang dilakukan dengan menggunakan metode atau algoritma K-Means dan dengan bantuan metode Elbow serta DBI untuk kasus Klasterisasi Kualitas Pendidikan Dasar Berdasarkan Elemen Ekosistem Pendidikan Menggunakan K-Means DBI di Kabupaten Magelang dapat disimpulkan bahwa:

1. Untuk memudahkan proses penentuan jumlah klaster optimal dengan menggunakan metode Elbow dapat menggunakan data yang sudah dinormalisasikan, yaitu data normalisasi berdasarkan elemen ekosistem pendidikan.
2. Data elemen ekosistem yang digunakan meliputi data sekolah, data guru dan data murid untuk setiap wilayah atau kecamatan di Kabupaten Magelang.
3. Klasterisasi yang dihasilkan berdasarkan nilai valid DBI, yaitu terdapat masing-masing 3 klaster untuk elemen sekolah, guru dan murid, dengan kategori tinggi untuk klaster1, sedang untuk klaster2 dan rendah untuk klaster3.
4. Hasil setiap klaster menunjukkan potensi dan tantangan bagi pihak terkait untuk ketersediaan jumlah elemen ekosistem pendidikan, sehingga diharapkan pihak yang berwenang dapat memberikan kebijakan demi berlangsungnya pemerataan kualitas pendidikan dasar di setiap wilayah.
5. Hasil untuk setiap elemen dalam setiap klaster terlampir pada tabel 14 dan 15.

Tabel 14. Hasil Klasterisasi untuk SD dan MI

SD			MI		
Klaster 1	Klaster 2	Klaster 3	Klaster 1	Klaster 2	Klaster 3
Muntilan	Salaman	Ngluwar	Kajoran	Salaman	Ngluwar
Mertoyudan	Borobudur	Salam	Kaliangkrik	Borobudur	Dukun
Grabag	Srumbung	Candimulyo	Secang	Salam	Muntilan
	Dukun	Kaliangkrik	Grabag	Srumbung	Sawangan
	Mungkid	Windusari		Mungkid	Candimulyo
	Sawangan	Ngablak		Mertoyudan	Tempuran
	Tempuran			Bandongan	Pakis
	Kajoran			Windusari	Ngablak
	Bandongan			Tegalrejo	
	Secang				
	Tegalrejo				
	Pakis				

Tabel 15. Hasil Klasterisasi untuk SMP dan MTs

SD			MI		
Klaster 1	Klaster 2	Klaster 3	Klaster 1	Klaster 2	Klaster 3
Salaman	Borobudur	Ngluwar	cluster1	cluster2	cluster3
Muntilan	Salam	Dukun	Salaman	Dukun	Borobudur
Mertoyudan	Srumbung	Candimulyo	Secang	Muntilan	Ngluwar
	Mungkid	Kajoran	Grabag	Mungkid	Salam
	Sawangan	Kaliangkrik		Kajoran	Srumbung
	Tempuran	Windusari		Kaliangkrik	Sawangan
	Bandongan	Ngablak		Bandongan	Candimulyo
	Secang			Windusari	Mertoyudan
	Tegalrejo			Tegalrejo	Tempuran
	Pakis				Pakis
	Grabag				Ngablak

DAFTAR PUSTAKA

- Apridayanti, A. Z., Fathurahman, M., & Prangga, S. (2024). Clustering of Province Based on Education Indicators Using K-Medoids. *Jurnal Varian*, 7(2), 10–16. <https://doi.org/10.30812/varian.v4i1.xxxxx>
- Hartanti, N. T., Seniwati, E., Pramitasari, R., Studi, P., Informasi, S., Amikom, U., Studi, P., Komputer, T., & Amikom, U. (2024). Metode Elbow K-Means dalam Implementasi Data Mining pada Pemetaan Penyebaran Guru SMK. *TEKNIKA*, 18(2), 501–512.
- Irwan, I., Sanusi, W., Anwar, A. S., & Rahman, A. (2023). The Implementation of Spatial Model with K-Means Clustering Method to Cluster Flood Affected Areas in Bone Regency. *ARRUS Journal of Social Sciences and Humanities*, 3(2), 186–195. <https://doi.org/10.35877/soshum1771>
- Juita, G., Jahera, J., Zulkifli, A., & Ananda, R. (2024). Standar Sarana dan Prasarana Pendidikan Dasar Dalam Meningkatkan Proses Pembelajaran Yang Efektif. *Elementary School*, 11, 572–582.
- Liu, R. (2022). Data Analysis of Educational Evaluation Using K-Means Clustering Method. *Computational Intelligence and Neuroscience*. <https://doi.org/10.1155/2022/3762431>
- Luchia, N. T., Handayani, H., Hamdi, F. S., Erlangga, D., & Octavia, S. F. (2022). Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 2(2), 35–41. <https://doi.org/10.57152/malcom.v2i2.422>
- Nurahman, N., & Aulia, D. D. (2023). Klasterisasi Pendidikan Masyarakat Untuk Mengetahui Daerah Dengan Pendidikan Terendah Menggunakan Algoritma K-Means. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 5(1), 38. <https://doi.org/10.36499/jinrpl.v5i1.7510>
- Oktavianto, H., & Fitriyah, N. Q. (2021). Implementasi Teknik Clustering Berbasis Validasi Cluster Pada Data Stunting. *Jurnal Teknologi Informasi (JUTECH)*, 2(2), 61–71. <https://doi.org/10.32546/jutech.v2i2.1778>
- Permadi, V. A., Tahalea, S. P., & Agusdin, R. P. (2023). K-Means and Elbow Method for Cluster Analysis of Elementary School Data. *Progres Pendidikan*, 4(1), 50–57. <https://doi.org/10.29303/prospek.v4i1.328>
- Rahmawati, O., & Fauzan, A. (2024). PROVINCIAL CLUSTERING BASED ON EDUCATION INDICATORS: K-MEDOIDS APPLICATION AND K-MEDOIDS OUTLIER HANDLING.

BAREKENG: Jurnal Ilmu Matematika Dan Terapan, 18(2), 1167–1178.
<https://doi.org/10.30598/barekengvol18iss2pp1167-1178>

- Syahrani, W., Studi Ilmu Komputer, P., & Sains Dan Teknologi, F. (2024). Penerapan Algoritma k-Means pada Clustering Data Penduduk Miskin untuk Penghapusan Kemiskinan Ekstrem. *Sistemasi: Jurnal Sistem Informasi*, 13(4), 2540–9719. <http://sistemasi.ftik.unisi.ac.id>
- Tu, K., Chen, Z., & Yue, M. (2024). A Max-Min-Max Algorithm for Large-Scale Robust Optimization. *Arxiv*, 11–47. <https://arxiv.org/abs/2404.05377>
- Yanto, D., Susanto, H., Zulkifli, K., & Gupron, F. R. (2023). Penerapan Algoritma K-Means Dalam Menentukan Kualitas Satuan Pendidikan Berdasarkan Nilai Internal Dan Eksternal. *JEECOM Journal of Electrical Engineering and Computer*, 5(2), 319–328. <https://doi.org/10.33650/jeecom.v5i2.6946>