

Implementasi Algoritma K-Means Dengan Normalisasi Min-Max Pada Analisis Data Ketidakbersekolahan Anak

Elsahday Tambunan¹, Yuni Br Limbeng², Sardo Sipayung³

^{1,2,3}Teknik Informatika, Universitas Katolik Santo Thomas, Indonesia

Email: ¹elsahdaytambunan@gmail.com, ²yunibrlimbeng@gmail.com, ³pinarsiphom@gmail.com

ABSTRAK

Anak-anak yang tidak bersekolah merupakan suatu masalah dalam dunia pendidikan yang masih menjadi tantangan, terutama di kalangan masyarakat dengan ekonomi rendah. Tingginya jumlah anak yang tidak mengenyam pendidikan dapat mengurangi kualitas sumber daya manusia dan memperbesar kesenjangan sosial. Penelitian ini bertujuan untuk mengkaji ketidakbersekolahan pada anak berdasarkan level pendidikan dan kelompok pengeluaran, dengan menggunakan pendekatan *data mining*. Metode yang diterapkan mencakup normalisasi Min-Max sebagai langkah awal dalam memproses data serta algoritma K-means Clustering untuk proses pengelompokan. Normalisasi Min-Max digunakan untuk menyamakan skala data dalam rentang 0 hingga 1, sehingga setiap variabel memiliki peran yang seimbang dalam perhitungan jarak. Data yang digunakan adalah data angka anak tidak sekolah Tahun 2023, yang mencakup tingkat pendidikan SD, SMP, dan SMA rentang kelompok pengeluaran dari kuantil 1 hingga 5. Temuan penelitian ini menunjukkan bahwa algoritma K-Means dengan $k = 3$ dapat mengelompokkan data menjadi tiga kluster utama, yakni tingkat ketidakbersekolahan yang tinggi, sedang, rendah. Ini mengindikasikan adanya hubungan antara level pengeluaran dan partisipasi anak dalam pendidikan.

Kata Kunci: Data Mining, K-Means Clustering, Normalisasi Min-Max, Ketidakbersekolahan Anak

ABSTRACT

Children who do not attend school is a problem in the world of education that remains a challenge, especially among low-income communities. The high number of children who do not receive education can reduce the quality of human resources and widen social inequality. This study aims to examine children's schoollessness based on education level and expenditure group, using a data mining approach. The methods applied include Min-Max normalization as the initial step in data processing and the K-means Clustering algorithm for the grouping process. Min-Max normalization is used to equalize the data scale in the range of 0 to 1, so that each variable has a balanced role in calculating the distance. The data used is the number of children out of school in 2023, which includes elementary, junior high, and high school education levels as well as expenditure groups from quantiles 1 to 5. The findings of this study indicate that the K-means algorithm with $k = 3$ can group the data into three main clusters, namely high, medium, and low levels of schoollessness. This indicates a relationship between expenditure levels and children's participation in education.

Keywords: Data Mining, K-Means Clustering, Min-Max Normalization, Child Unschooling

Penulis Korespondensi:

Elsahday Tambunan

Email: elsahdaytambunan@gmail.com

Article Info

Diterima: 24 Januari 2026

Direvisi: 29 Januari 2026

Disetujui: 30 Januari 2026

This is an open access article under the [CC BY](#) license.



1. PENDAHULUAN

Pendidikan adalah dasar penting untuk perkembangan suatu negara. Melalui investasi di bidang pendidikan, diharapkan akan tercipta sumber daya manusia yang berkualitas, inovatif, dan mempunyai daya saing yang tinggi [1]. Namun, di Indonesia, masalah anak yang tidak bersekolah tetap menjadi tantangan yang serius yang menghambat kemajuan. Data dari Badan Pusat

Statistik (BPS) dengan jelas menunjukkan bahwa jumlah anak yang tidak bersekolah tetap tinggi, terutama pada jenjang pendidikan menengah dan di daerah-daerah dengan kondisi sosial-ekonomi yang kurang mendukung.

Anak-anak yang tidak bersekolah adalah sebuah masalah yang memiliki banyak aspek dan menghalangi kemajuan suatu negara. Di Indonesia, meskipun banyak usaha telah dilakukan, jumlah anak yang tidak bersekolah masih menjadi isu penting. Menurut data dari Badan Pusat Statistik (BPS), jumlah siswa di tingkat Sekolah Menengah Pertama (SMP) lebih banyak dari pada di Sekolah Dasar (SD). Ini menunjukkan bahwa ada tantangan dalam transisi pendidikan yang harus diatasi [2].

Alasan mengapa anak-anak tidak bersekolah sangat beragam dan rumit. Ini termasuk kondisi ekonomi keluarga, ketidaktahuan orang tua mengenai nilai pendidikan, serta faktor-faktor sosial dan budaya. Memahami faktor-faktor ini secara detail adalah hal yang krusial untuk mengembangkan kebijakan dan tindakan yang efektif [3]. Di Zaman digital saat ini, pengguna teknologi informasi untuk menganalisis data pendidikan terus meningkat. Metode K-Means, yang merupakan jenis *unsupervised learning*, dapat digunakan untuk mengelompokkan data berdasarkan sifat yang sama. Dengan menggunakan algoritma K-Means pada data anak yang tidak bersekolah, kita dapat menemukan kelompok-kelompok daerah atau orang-orang yang memiliki karakteristik ATS yang beragam [4].

Dalam upaya memahami fenomena ketidakbersekolahan anak, tidak cukup hanya mendeskripsikan faktor-faktor yang mempengaruhinya, tetapi juga diperlukan pendekatan analitis kuantitatif yang mampu menemukan pola dan hubungan antar variabel. Salah satu pendekatan yang sering digunakan dalam penelitian pendidikan adalah *clustering* atau pengelompokan data, yang merupakan bagian dari *unsupervised learning* dalam *machine learning* [5]. Algoritma K-means adalah salah satu teknik *clustering* yang paling populer karena kesederhanaan, kecepatan komputasi, dan kemampuannya dalam mengelompokkan data menjadi beberapa kelompok berdasarkan kemiripan karakteristiknya [6].

K-Means secara prinsip membagi N objek data menjadi K kelompok sehingga objek dalam satu kluster memiliki kemiripan yang lebih tinggi dibandingkan dengan objek pada kluster lain. Penerapan K-Means dalam penelitian pendidikan telah banyak dilakukan untuk berbagai tujuan, seperti mengidentifikasi kelompok sekolah berdasarkan kualitas pendidikan, menggambar pola prestasi siswa, atau mengelompokkan wilayah dengan karakteristik pendidikan yang berbeda [7].

Namun, sebelum dilakukan proses *clustering*, satu tahapan penting dalam pra-pemrosesan data adalah normalisasi data. Normalisasi Min-Max merupakan teknik yang umum digunakan yang menormalisasi nilai variabel numerik agar berada dalam rentang tertentu, biasanya antara 0 atau 1. Tujuan utama normalisasi adalah memastikan setiap fitur memiliki kontribusi yang seimbang dalam proses *clustering* agar variabel dengan rentang besar tidak mendominasi algoritma. Hal ini terbukti meningkatkan kualitas kluster yang dihasilkan, khususnya ketika dataset mengandung fitur numerik dengan skala yang sangat berbeda [6].

Dalam konteks analisis pendidikan, berbagai studi telah menunjukkan efektivitas pendekatan *clustering* untuk mengevaluasi data pendidikan [8]. Sebagai contoh, penelitian yang mengelompokkan sekolah menengah atas berdasarkan indikator akreditasi, jumlah guru, dan fasilitas menggunakan K-Means dengan normalisasi Min-Max, menemukan tiga kelompok sekolah yang berbeda yang dapat digunakan sebagai dasar kebijakan pengembangan pendidikan setempat.

Studi ini memiliki tujuan untuk menerapkan algoritma K-Means bersama dengan normalisasi Min-Max pada data mengenai anak-anak yang tidak bersekolah di Indonesia [9]. Melalui analisis ini, diharapkan faktor-faktor utama yang mempengaruhi ketidakbersekolahan dapat dikenali, serta menghasilkan saran kebijakan yang berbasis data guna meningkatkan partisipasi dalam pendidikan. Selain itu, penelitian ini diharapkan dapat berkontribusi pada kemajuan metodologi analisis data pendidikan yang lebih canggih dan efektif [10].

Untuk menangani masalah anak-anak yang tidak bersekolah, sangat penting untuk mengenali penyebab utamanya dan menemukan kelompok anak yang paling berisiko. Analisis data bisa sangat membantu dalam situasi ini dengan memberikan pemahaman yang lebih baik tentang pola-pola ketidakbersekolahan dan faktor risiko yang ada. Salah satu metode yang sering digunakan dalam analisis data adalah pengelompokan, yang memberi kesempatan kepada penelitian untuk mengelompokkan data berdasarkan ciri-ciri yang sama [11]. Metode ini membantu menghindari dominasi fitur dengan nilai lebih tinggi dalam proses pengelompokan dan memastikan bahwa semua fitur memberikan kontribusi yang seimbang terhadap hasil [12].

Sebelum menggunakan algoritma K-Means, sangat penting untuk menormalkan data agar semua fitur memiliki skala yang serupa. Salah satu metode Normalisasi yang sering digunakan adalah Normalisasi Min-Max, yang mengubah nilai fitur sehingga berada dalam rentang 0 hingga 1 [13].

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan cara deskriptif melalui teknik *data mining*, khususnya menggunakan K-Means untuk *Clustering*. Pilihan pendekatan kuantitatif diambil karena data yang digunakan merupakan angka yang diperoleh dari pengukuran statistik, sedangkan *clustering* digunakan untuk menemukan pola dan mengelompokkan data tanpa label kelas, yang dikenal sebagai pembelajaran tanpa pengawasan.

Data yang dipakai dalam studi ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Republik Indonesia, yaitu informasi mengenai angka anak tidak sekolah menurut jenjang pendidikan dan kelompok pengeluaran. Informasi ini meliputi persentase anak yang tidak bersekolah pada level SD, SMP, dan SMA yang dikelompokkan ke dalam lima kategori pengeluaran.

2.1 Tahapan Penelitian

Tahapan penelitian dilakukan secara sistematis sebagai berikut:

(1) Pengumpulan Data

Mengunduh dataset BPS berupa angka anak tidak bersekolah menurut jenjang dan pengeluaran.

(2) Pra-Pemrosesan Data

Data diubah ke skala 0-1 agar setiap variabel memiliki kontribusi seimbang dalam proses klusterisasi [14].

(3) Proses *Clustering* K-Means

Data yang sudah dinormalisasi dikelompokkan menggunakan algoritma K-Means. Nilai K ditentukan melalui metode evaluasi.

(4) Analisis dan Interpretasi hasil *clustering*

Menganalisis dan menyimpulkan temuan dan merekomendasikan implikasi kebijakan berdasarkan hasil *clustering*.

Tahapan penelitian berbasis *data mining* ini mengacu pada tahapan penelitian yang digunakan oleh Irwansyah dkk. (2023) dalam analisis data pendidikan menggunakan algoritma K-Means [7].

2.2 Normalisasi Min-Max

Normalisasi Min-Max bertujuan untuk mentransformasikan setiap variabel numerik kendala rentang antara 0 sampai 1, agar variabel dengan nilai besar tidak mendominasi proses *clustering* [15]. Rumus Normalisasi Min-Max:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

dengan:

X' = nilai setelah dinormalisasi

X = nilai asli

X_{min} = nilai minimum variabel

X_{max} = nilai maksimum variabel

2.3 Algoritma K-Means

K-Means adalah algoritma pengelompokan (*clustering*) yang bekerja dengan membagi dataset ke dalam K kelompok berdasarkan kemiripan (*similarity*) jarak menggunakan *centroid*. K-Means banyak digunakan dalam analisis data sosial dan pendidikan [16]. Langkah kerja K-Means:

1. Tentukan jumlah kluster K .
2. Inisialisasi *centroid* awal secara acak.
3. Hitung jarak tiap titik data ke setiap *centroid*.
4. Tempatkan setiap data ke kluster dengan *centroid* terdekat.
5. Hitung ulang *centroid* tiap kluster.
6. Ulangi langkah 3–5 hingga *centroid* stabil.

Euclidean Distance:

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

3. HASIL DAN PEMBAHASAN

3.1 Data Asli Angka Anak Tidak Sekolah Menurut Jenjang Pendidikan dan Kelompok Pengeluaran

Data yang digunakan untuk penelitian ini diperoleh dari Badan Pusat Statistik (BPS) dengan judul Angka Anak Tidak Sekolah Berdasarkan Jenjang pendidikan dan Kelompok Pengeluaran Tahun 2023. Informasi ini menunjukkan persentase anak yang tidak bersekolah pada tiga level pendidikan, yaitu SD/ sederajat, SMP/ sederajat, dan SMA/ sederajat, yang dikelompokkan menurut lima kategori pengeluaran.

Pembagian kelompok pengeluaran dilakukan berdasarkan tingkat kesejahteraan ekonomi rumah tangga, dimana Kuintil 1 mewakili kelompok dengan pengeluaran terendah dan Kuintil 5 mewakili kelompok dengan pengeluaran tertinggi. Penelitian ini melibatkan tiga atribut numerik (SD, SMP, SMA) dan lima objek data (kuintil).

Awal hasil menunjukkan bahwa Kuintil 1 memiliki angka anak tidak sekolah tertinggi di semua level pendidikan, terutama pada tingkat SMA, sedangkan Kuintil 5 mencatat nilai terendah. Temuan ini menunjukkan adanya hubungan yang kuat antara kondisi ekonomi dan partisipasi dalam pendidikan.

Tabel 1. Data Asli

Kelompok Pengeluaran	Angka Anak Tidak Sekolah Menurut Jenjang Pendidikan dan Kelompok Pengeluaran		
	SD / Sederajat 2023	SMP / Sederajat 2023	SMA / Sederajat 2023
Kuintil 1	1,15	10,11	30,66
Kuintil 2	0,58	7,33	23,21
Kuintil 3	0,58	6,14	19,65
Kuintil 4	0,46	5,37	19,31
Kuintil 5	0,38	5,05	14,67

3.2 Hasil Normalisasi Min-Max

Sebelum *clustering* dilaksanakan, data dinormalisasi dengan metode Min-Max Normalization yang memiliki rentang nilai dari 0 sampai 1. Tujuan dari normalisasi ini adalah untuk menyamakan skala antara variabel-variabel sehingga satu variabel tidak menguasai variabel lainnya dalam perhitungan jarak Euclidean.

Hasil normalisasi Min-Max menunjukkan bahwa:

- Kuintil 1 memiliki nilai normalisasi tertinggi (1,00) pada seluruh jenjang pendidikan.
- Kuintil 5 memiliki nilai normalisasi terendah (0,00).
- Kuintil 2 hingga Kuintil 4 berada pada rentang nilai menengah.

Normalisasi ini membuktikan bahwa data telah siap digunakan dalam proses K-Means Clustering karena setiap atribut memiliki kontribusi yang seimbang.

Tabel 2. Hasil Normalisasi Min-Max

HASIL NORMALISASI MIN - MAX			
KELOMPOK PENGELUARAN	SD	SMP	SMA
KUINTIL 1	1,00	1	1
KUINTIL 2	0,26	0,451	0,534
KUINTIL 3	0,26	0,215	0,311
KUINTIL 4	0,10	0,063	0,290
KUINTIL 5	0,00	0	0

3.3 Penemuan Jumlah Kluster (K)

Jumlah kluster ditentukan sebanyak $K = 3$, dengan pertimbangan tujuan analisis untuk mengelompokkan tingkat ketidakbersekolahan menjadi tiga kategori utama, yaitu:

- Kluster 1 (C1) : Tingkat Anak Tidak Sekolah Tinggi
- Kluster 2 (C2) : Tingkat Anak Tidak Sekolah Sedang
- Kluster 3 (C3) : Tingkat Anak Tidak Sekolah Rendah

Penentuan ini sesuai dengan kebutuhan interpretasi kebijakan pendidikan dan memudahkan pemangku kepentingan dalam memahami hasil pengelompokan.

3.4 Proses K-Means Clustering

3.4.1 Inisialisasi Centroid Awal

Centroid awal ditentukan berdasarkan data hasil normalisasi:

- C1 diambil dari nilai tertinggi (Kuintil 1),
- C2 diambil dari nilai menengah (Kuintil 2),
- C3 diambil dari nilai terendah (Kuintil 5).

Tabel 3. K-Means Clustering

K-MEANS CLUSTERING			
KELOMPOK PENGELUARAN	SD	SMP	SMA
KUINTIL 1	1,00	1	1
KUINTIL 2	0,260	0,451	0,534
KUINTIL 3	0,260	0,215	0,311
KUINTIL 4	0,104	0,063	0,290
KUINTIL 5	0	0	0
JUMLAH VARIABEL		3	
JUMLAH DATA		5	

3.4.2 Perhitungan Jarak Euclidean

Pada setiap iterasi, jarak antara data dan centroid dihitung menggunakan Euclidean Distance. Data kemudian dikelompokkan ke dalam kluster dengan jarak terdekat.

3.4.3 Iterasi dan Pembaruan Centroid

Proses iterasi dilakukan hingga centroid tidak mengalami perubahan signifikan. Berdasarkan hasil perhitungan:

- Iterasi berhenti pada Iterasi kel-4, karena tidak terjadi perubahan keanggotaan kluster.
- Hal ini menandakan bahwa algoritma telah mencapai kondisi konvergen.

Tabel 4. Iterasi Satu

ITERRASI 1			
CLUSTER	SD	SMP	SMA
C1	1,00	1	1
C2	0,260	0,451	0,534
C3	0	0	0

Tabel 5. Hasil perhitungan Jarak Iterasi Satu

HASIL PERHITUNGAN JARAK (ITERASI 1)				
KELOMPOK	d (C1)	d(C2)	d(C3)	Kluster
KUINTIL 1	0	1,033	1,732	C1
KUINTIL 2	1,0329	0	0,745	C2
KUINTIL 3	1,2797	0,324	0,459	C2
KUINTIL 4	1,4780	0,484	0,315	C3
KUINTIL 5	1,7321	0,745	0	C3

Tabel 6. Centroid Baru

CENTROID BARU Iterasi 2			
CLUSTER	SD	SMP	SMA
C1	0,00	1,03	1,73
C2	1,156	0,162	0,602
C3	1,605	0,615	0,157

Tabel 7. Hasil Perhitungan Jaraj Iterasi Dua

HASIL PERHITUNGAN JARAK (ITERASI 2)				
KUINTIL 1	1,2398	0,9407	1,1067	C2
KUINTIL 2	1,3571	0,9444	1,4066	C2
KUINTIL 3	1,6595	0,9441	1,4117	C2
KUINTIL 4	1,7407	1,1022	1,6046	C2
KUINTIL 5	2,0167	1,3138	1,7258	C2

Tabel 8. Centroid Baru Iterasi Tiga

ITERRASI 3			
CLUSTER	SD	SMP	SMA
C1	0,00	1,23	1,73
C2	0,000	0,000	0,602
C3	0,000	0,000	0,000

Tabel 9. Centroid Iterasi Empat

ITERRASI 4			
CLUSTER	SD	SMP	SMA
C1	0,00	0,00	0,00
C2	0,000	0,000	0,000
C3	0,000	0,000	0,000

3.4.4 Hasil Akhir Clustering

Berdasarkan hasil akhir K-Means Clustering, diperoleh pengelompokan sebagai berikut:

Tabel 10. Hasil Akhir K-Means Clustering

KELOMPOK PENGELUARAN	CLUSTER	INTERPRETASI
KUINTIL 1	C1	Tingkat Anak Tidak Sekolah Tinggi
KUINTIL 2	C1,C2	Tingkat Anak Tidak Sekolah Tinggi, Tingkat Sedang
KUINTIL 3	C2	Tingkat Sedang
KUINTIL 4	C3	Tingkat Rendah
KUINTIL 5	C3	Tingkat Rendah

3.5 Pembahasan

Hasil dari pengelompokan memberikan dukungan pada penemuan bahwa aspek ekonomi memiliki dampak yang besar terhadap anak yang tidak bersekolah. Kuintil 1, yang termasuk dalam kelompok dengan tingkat tinggi, menunjukkan bahwa keterbatasan finansial menjadi salah satu penghalang utama untuk mendapatkan pendidikan, terutama di tingkat menengah dan atas.

Di sisi lain, Kuintil 4 dan Kuintil 5 berada dalam kelompok dengan tingkat rendah, yang menunjukkan bahwa kondisi ekonomi yang lebih baik memberikan kesempatan lebih besar bagi anak untuk melanjutkan pendidikan. Penggunaan algoritma K-Means dengan normalisasi Min-Max terbukti efektif dalam menyortir pola ketidakbersekolahan anak berdasarkan data yang objektif.

4. KESIMPULAN

Berdasarkan analisis data mengenai anak-anak yang tidak bersekolah menurut jenjang pendidikan dan kategori pengeluaran tahun 2023, yang dilakukan dengan algoritma K-Means serta normalisasi Min-Max, disimpulkan bahwa metode *data mining* ini dapat dengan objektif dan sistematis mengelompokkan tingkat ketidakbersekolahan anak. Normalisasi Min-Max berhasil menyamakan skala data untuk jenjang SD, SMP, dan SMA sehingga setiap variabel berkontribusi seimbang dalam perhitungan jarak Euclidean selama proses pengelompokan.

Analisis pengelompokan mengungkapkan terdapat tiga kelompok utama, yaitu tingkat ketidakbersekolahan yang tinggi, sedang, dan rendah. Kategori Kuintil 1 secara konsisten termasuk dalam kluster dengan tingkat ketidakbersekolahan yang tinggi, menunjukkan bahwa kelompok ekonomi terendah menghadapi risiko lebih besar terhadap putus sekolah di semua jenjang pendidikan, terutama di jenjang SMA. Di sisi lain, Kuintil 2 dan Kuintil 3 berada di kluster tingkat sedang, yang mencerminkan adanya peningkatan partisipasi pendidikan seiring dengan perbaikan kemampuan ekonomi walaupun masih ada risiko ketidakbersekolahan. Kuintil 4 dan Kuintil 5 dikelompokkan ke dalam kluster tingkat rendah, yang menunjukkan bahwa mereka yang memiliki pengeluaran lebih tinggi mendapatkan akses pendidikan yang lebih baik dan mengalami angka ketidakbersekolahan yang rendah.

Dengan demikian, studi ini menunjukkan bahwa keadaan ekonomi keluarga memiliki dampak besar terhadap tingkat ketidakbersekolahan anak di Indonesia. Penggunaan algoritma K-Means dengan normalisasi Min-Max tidak hanya berhasil dalam mengidentifikasi pola ketidakbersekolahan anak, tetapi juga dapat berfungsi sebagai dasar analisis yang kuat untuk merancang kebijakan pendidikan yang lebih efektif, terutama dalam usaha mengurangi jumlah anak tidak bersekolah di kalangan masyarakat berpenghasilan rendah.

UCAPAN TERIMA KASIH

Penulis mengucapkan puji dan syukur ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga penelitian yang berjudul “Implementasi Algoritma K-Means dengan Normalisasi Min-Max pada Analisis Data Ketidakbersekolahan Anak” dapat diselesaikan dengan baik.

Penulis juga mengucapkan terima kasih kepada Badan Pusat Statistik (BPS) yang telah menyediakan data resmi dan terbuka mengenai angka anak tidak sekolah di Indonesia, sehingga penelitian ini dapat dilaksanakan secara objektif dan berbasis data yang valid.

Penulis menyampaikan terima kasih yang sebesar-besarnya kepada Universitas Katolik Santo Thomas sebagai institusi pendidikan yang telah memberikan fasilitas, lingkungan akademik, serta dukungan dalam proses penyelesaian penelitian ini. Ucapan terima kasih juga penulis sampaikan kepada dosen pembimbing dan seluruh dosen Universitas Katolik Santo Thomas yang telah memberikan bimbingan, arahan, serta ilmu pengetahuan sehingga penelitian ini dapat diselesaikan dan disajikan dalam bentuk karya ilmiah.

REFERENSI

- [1] Sarfa Wassahua, “Journal MANAGER,” *J. Pendidik.*, vol. 1, no. 2, hal. 93–113, 2016.
- [2] A. Dan, V. Data, A. Tidak, S. Di, I. menurut, dan P. Bi, “1 Sistem Informasi, Fakultas Teknik dan Ilmu Komputer , Universitas Islam Indragiri,” vol. 3, no. 1, hal. 28–36, 2025.
- [3] A. Maroddin, I. Muh. Said, dan Kasmawati, “Social landscape journal,” *Soc. Landsc. J.*, vol. 3, no. 2, hal. 9–18, 2022.
- [4] N. Hayati dan R. Supriadi, “ANALISIS FAKTOR PENYEBAB ANAK TIDAK MELANJUTKAN PELNDIDIKAN KE SEKOLAH MENENGAH ATAS Nur Hayati, Supriadi, Rustiyarso Program Studi pendidikan Sosiologi FKIP UNTAN Pontianak,” no. 1, hal. 1–11, 2018.
- [5] S. A. Fahmi, A. Ikhwan, dan F. H. Sibarani, “Implementasi Data Mining Dengan Menggunakan Metode K-Means Clustering Untuk Menentukan Penjualan Sparepart Mobil,” vol. 4, no. 1, hal. 204–209, 2025.
- [6] T. A. Siagian, N. Nurdin, dan M. Ula, “Clustering of the BLElst Senior High Schools in Serdang Bedagai Regency Using the K-Means Method,” *J. Sist. Komput. dan Inform.*, vol. 6, no. 4, hal. 202–208, 2025, doi: 10.30865/json.v6i4.8669.
- [7] M. A. Irwansyah, YL. N. Alinda, R. N. Aini, dan I. A. Alftrah, “KLASIFIKASI KUALITAS PENDIDIKAN SEKOLAH DASAR,” vol. 7, no. 1, hal. 1008–1020, 2025.
- [8] R. Sari, M. Yasin, dan U. Asahan, “Penerapan Data mining untuk Clustering Kondisi Sosial Ekonomi Berdasarkan Kepemilikan Jaminan Kesehatan Menggunakan Algoritma untuk merumuskan kebijakan yang lebih efektif. Setelah kelompok-kelompok masyarakat,” no. September, 2025.
- [9] D. Anggraini, P. Korespondensi, dan P. Mesin, “Data Mining Pendidikan : Prediksi Gaya Belajar Mahasiswa Educational Data Mining : Predicting of Engineering Students ’ Learning Styles Using Machine Learning,” vol. 11, no. 3, hal. 563–

- 572, 2025.
- [10] F. K. Oktorina, T. Rekeyasa, L. Politeknik, T. Informatika, dan P. Kampar, "Penerapan Data Mining Untuk Analisis Hasil Belajar Mahasiswa (Studi Kasus Mahasiswa Teknik Informatika Politeknik Kampar)," vol. 5, no. 2, 2025.
 - [11] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *J. Ilm. Inform. dan Ilmu Komput.*, vol. 3, no. 1, hal. 46–56, 2024, doi: 10.58602/jima-ilkom.v3i1.26.
 - [12] Faozan Dwiki Ramadana, Wahyu Putra Pratama, Cannels Lingga Yogario, Abdul Khohar, dan Ito Setiawan, "Implementasi Algoritma K-Means Clustering terhadap Tingkat Kepuasan Peserta LKP Multi Talenta Komputer Purwokerto," *Mars J. Tek. Mesin, Ind. Elektro Dan Ilmu Komput.*, vol. 3, no. 1, hal. 184–193, 2025, doi: 10.61132/mars.v3i1.675.
 - [13] YL. K. Jain dan S. K. Bhandarel, "International Journal of Computer and Communication Technology Min Max Normalization Based Data Perturbation Method for Privacy Protection," vol. 4, no. 4, 2013, doi: 10.47893/IJCCT.2013.1201.
 - [14] Maghfiroh, S., Fatah, Z. "Analisis Data Mining dengan Algoritma K-Means Clustering untuk Menentukan Siswa Berprestasi di MTS miftahul ulum bengkak," vol. 4, 2025.
 - [15] S. N. Sinha dan R. L. Yadav, "Analysis of Data using K-Means Clustering Algorithm with Min Max Function," vol. 58, no. 2, hal. 82–84, 2018.
 - [16] F. Dellia *et al.*, "Penerapan Algoritma K-Means Clustering Dalam Pengelompokan Kepadatan Penduduk Application of K-Means Clustering Algorithm in Population Density," vol. 9, no. 3, hal. 373–386, 2025.