

Analisis Kelompok Lansia Berdasarkan Kategori Usia Dengan Metode K-Means Clustering

Lela Lailatul Kaamilah

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon

Mulyawan

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon

Alamat: Cirebon, Jawa Barat, Indonesia 45131

Korespondensi penulis: lelakamilah1@gmail.com

Abstract.

The application of k-means can be used to group the number of elderly people based on age category. By using this algorithm, groups of elderly residents who have the same age characteristics can be determined. Elderly (elderly) is the share of the population aged 60 years and over, the health department categorizes the elderly based on their age level, namely: early elderly, late elderly, and elderly). The increasing number of elderly people worldwide has created problems in managing the welfare of the elderly population. Inaccurate information regarding the distribution of the number of elderly people in each age category makes it difficult to make the right decisions to improve the welfare of the elderly population. In this case, the application of K-Means clustering can be used as a tool in grouping the number of elderly people based on age category. With this algorithm, data can be grouped quickly and efficiently, so that it can assist in making appropriate decisions to improve the welfare of the elderly population. However, the K-Means Clustering algorithm is only used as a tool, and must be strengthened with proper analysis and recommendations. In the village of Cimari, Cikoneng District, Ciamis Regency, in terms of grouping the elderly based on age categories, there are limitations in managing data on the elderly population, namely by manual method which requires quite a long time. The design method used is data collection: the number of elderly people and their age category. This data was obtained from data from Cimari villagers Based on the grouping results using K-Means Clustering on the grouping of the elderly population by age category, the Davies Bouldin results were 0.263, cluster 0 contained 119 items, cluster 1 contained 101 items, and cluster 2 contained 80 items.

Keywords: Data mining, Clustering, K-Means Clustering, elderly people.

Abstrak.

Penerapan k-means dapat digunakan untuk melakukan pengelompokan jumlah penduduk lansia berdasarkan kategori usia. Dengan menggunakan algoritma ini, dapat ditentukan kelompok-kelompok penduduk lansia yang memiliki karakteristik usia yang sama. Lanjut Usia (lansia) merupakan bagian penduduk berusia 60 tahun ke atas, departemen kesehatan mengkategorikan usia lanjut berdasarkan tingkatan usianya yaitu : lansia awal, lansia akhir, dan manula. Pertumbuhan jumlah penduduk lansia yang

semakin meningkat di seluruh dunia menimbulkan masalah dalam pengelolaan kesejahteraan penduduk lansia. Informasi yang tidak akurat mengenai distribusi jumlah penduduk lansia dalam setiap kategori usia menyulitkan dalam mengambil keputusan yang tepat untuk meningkatkan kesejahteraan penduduk lansia. Dalam hal ini, penerapan K-Means clustering dapat digunakan sebagai alat bantu dalam pengelompokan jumlah penduduk lansia berdasarkan kategori usia. Dengan algoritma ini, data dapat dikelompokkan dengan cepat dan efisien, sehingga dapat membantu dalam pengambilan keputusan yang sesuai untuk meningkatkan kesejahteraan penduduk lansia. Namun, algoritma K-Means Clustering hanya digunakan sebagai alat bantu saja, dan harus diperkuat dengan analisis dan rekomendasi yang tepat. Di desa Cimari Kecamatan Cikoneng Kabupaten Ciamis dalam hal pengelompokan lansia berdasarkan kategori usia ada keterbatasan dalam mengelola data penduduk lansia yaitu dengan cara manual yang memerlukan waktu cukup lama. Metode perancangan yang diterapkan yaitu pengumpulan data: jumlah penduduk lansia dan kategori usianya. Data ini diperoleh dari data penduduk desa Cimari. Berdasarkan hasil pengelompokan menggunakan K-Means Clustering terhadap pengelompokan jumlah penduduk lansia berdasarkan kategori usia diperoleh hasil Davies Bouldin sebesar 0.263, cluster 0 terdapat 119 item, cluster 1 terdapat 101 item, dan cluster 2 terdapat 80 item.

Kata kunci : Data mining, Clustering, K-Means Clustering, penduduk lansia

LATAR BELAKANG

Lansia merupakan kelompok berkebutuhan khusus dan memerlukan perhatian khusus dari pemerintah dan organisasi sosial. Namun, perencanaan dan pengelolaan program yang tepat seringkali sulit dilakukan karena terbatasnya data yang tersedia. Oleh karena itu, penerapan k-means dapat digunakan untuk mengelompokkan jumlah lansia ke dalam kategori usia untuk membantu dalam perencanaan dan penyelenggaraan program yang tepat. Pengolahan Jumlah penduduk dalam penelitian ini dikelompokkan berdasarkan usia menurut departemen Kesehatan yang mengkategorikan usia lanjut berdasarkan tingkatan usianya, yaitu : lansia awal, lansia akhir, dan manula yang diperoleh data jumlah penduduk per dusun di Desa Cimari Kecamatan Cikoneng Kabupaten Ciamis. Melalui data ini peneliti akan terfokus pada pengelompokkan jumlah penduduk lansia berdasarkan kategori usia Berdasarkan uraian diatas masalah yang ditemukan yaitu Di desa Cimari dalam hal pengelompokan lansia berdasarkan kategori usia ada keterbatasan yaitu dengan cara manual yang memerlukan waktu cukup lama.

Tujuan dari penelitian ini adalah untuk mengelompokkan jumlah lansia berdasarkan kategori usia menurut departemen kesehatan sehingga dapat digunakan untuk analisis lebih lanjut seperti perencanaan program pemerintah untuk penduduk lanjut usia.

Berdasarkan penjelasan diatas, maka penulis tertarik untuk mengambil judul Analisis Kelompok Lansia Berdasarkan Kategori Usia Dengan Metode K-Means Clustering dimana data penduduk lansia diperoleh dari data penduduk Desa Cimari Kecamatan Cikoneng Kabupaten Ciamis. Menggunakan metode algoritma k-means clustering karena merupakan metode yang tepat untuk digunakan dalam mengelompokkan data penduduk lansia berdasarkan kategori usia

KAJIAN TEORITIS

A. Data Mining

Menemukan pengetahuan tersembunyi dalam basis data dikenal sebagai data mining. Data mining adalah proses semi-otomatis yang mengekstraksi dan mengidentifikasi bahan pengetahuan yang dapat digunakan dan berpotensi bermanfaat dari basis data besar menggunakan pendekatan statistik, matematika, kecerdasan buatan, dan pembelajaran mesin. (Jurnal et al., 2021)

B. Clustering

Clustering adalah teknik data mining tanpa pengawasan. Artinya bahwa karakteristik setiap cluster tidak ditentukan sebelumnya, tetapi didasarkan pada kesamaan atribut dari kelompok atau cluster tersebut. Berdasarkan kesamaan atribut antara set data yang berbeda, cluster mengatur data ke dalam kelompok atau cluster. Dalam ruang multidimensi, karakteristik yang sebanding ini ditampilkan sebagai titik. Dua jenis metode pengelompokan yang digunakan dalam data mining adalah *Hierarchical Clustering* dan *non-hierarchical Clustering*. (Andini & Farokhah, 2022)

C. K-Means Clustering

Metode k-means clustering memisahkan kumpulan data menjadi sejumlah cluster yang telah ditentukan melalui pengelompokan data berulang. Metode pengelompokan k-means adalah teknik yang banyak digunakan untuk menilai kinerja kluster karena mudah diatur, dijalankan, dan diubah. (Homepage et al., 2021)

D. RapidMiner

RapidMiner adalah perangkat lunak *open source* yang tersedia untuk semua orang. RapidMiner adalah alat untuk analisis data yang menggunakan berbagai metode, seperti metode deskriptif dan prediktif. Java adalah bahasa pemrograman yang digunakan oleh RapidMiner. Perangkat lunak RapidMiner dapat digunakan untuk memproses penambahan data berkat teknologi algoritma komputasi mutakhir dan analisis data berbasis komputer. (Ratna Sari et al., 2020)

E. Microsoft Office Excel

Microsoft Office Excel adalah program aplikasi lembar kerja spreadsheet yang dibuat dan didistribusikan oleh Microsoft Corporation untuk sistem operasi Microsoft Windows dan Mac.OS. Aplikasi ini memiliki kemampuan kalkulasi dan grafik dengan menggunakan strategi pemasaran Microsoft yang agresif telah menjadikan Microsoft Excel salah satu program komputer paling populer yang pernah digunakan di computer mikro. (Yulianti et al., 2019)

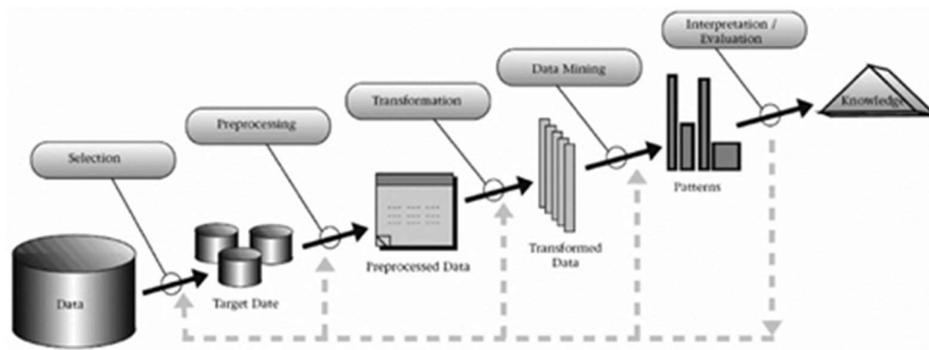
F. Kategori Usia

Menurut Departemen Kesehatan Republik Indonesia tahun 2009 kategori usia, sebagai berikut : Balita usia 0 – 5, Kanak-kanak usia 5 – 11, Remaja awal usia 12 – 16 , Remaja akhir usia 17 – 25, Dewasa awal usia 26 – 35, Dewasa akhir usia 36 – 45, Lansia awal usia 46 – 55, Lansia akhir usia 56 – 65, Manula usia 65 – ke atas(Sonang et al., 2019)

G. Lanjut Usia

Lanjut usia berada pada tahap akhir dari proses penuaan. Penuaan merupakan satu kondisi yang terjadi dalam kehidupan manusia. Proses menua merupakan proses seumur hidup, yang tidak hanya dimulai pada waktu tertentu, tetapi sejak awal kehidupan. Penuaan merupakan proses alami, artinya seseorang telah melewati tiga tahap yaitu masa kanak-kanak, dewasa, dan tua. Proses penuaan adalah proses yang berkaitan dengan usia.. (Dian et al., 2021)

METODE PENELITIAN



Gambar 1 Tahapan KDD

Keterangan gambar 1 :

1. Seleksi Data (*Selection*)

Sebelum tahap Knowledge Discovery Database (KDD) dari penemuan informasi dapat dimulai, pemilihan data dari kumpulan data operasional harus diselesaikan. Data operasional database operasional disimpan dalam file yang berbeda dari data yang dipilih untuk digunakan dalam proses data mining.

2. Pemilihan Data (*Preprocessing*)

Proses Preprocessing meliputi penghapusan data duplikat, pengecekan data yang tidak konsisten, dan perbaikan kesalahan pada data. Juga dilakukan proses enrichment, artinya proses “memperkaya” data yang ada dengan data atau informasi eksternal

3. Transformasi Data (*Data Transformation*)

Pada tahap ini ialah melakukan transformasi data yang belum memiliki entitas jelas untuk diubah menjadi bentuk data yang valid atau siap menjalani proses Data Mining..

4. Data mining

Data mining adalah proses menjelajahi data yang dipilih untuk pola dan informasi yang menarik menggunakan teknik dan prosedur tertentu..

5. *Interpretation/ Evaluation*

Evaluasi ialah tahap terakhir, yang dilakukan dalam tahap ini yaitu pembentukan keluaran yang nantinya mudah dimengerti yang bersumber pada proses data mining pola informasi.

Metode pengumpulan data

Sumber data dalam penelitian adalah dari Desa Cimari. Data ini meliputi data penduduk lansia Desa Cimari tahun 2021, waktu pengambilan data dilaksanakan pada tanggal 6 Januari 2023. Data penelitian ini termasuk ke dalam data sekunder karena data diperoleh dari laporan pemerintah desa.

Teknik atau cara yang dilakukan untuk mengumpulkan data ini yaitu observasi karena datang langsung ke kantor Desa Cimari. Jenis teknik pengumpulan data yang sering digunakan Wawancara, Observasi, dokumentasi. Wawancara merupakan salah satu teknik yang dilakukan untuk mengumpulandata atau informasi yang berkaitan dengan masalah penelitian di lapangan. Dengan wawancara peneliti dapat mengetahui data dan informasi yang dibutuhkan Penulis mengajukan pertanyaan kepada narasumber untuk mengumpulkan dataatau informasi yang diperlukan dalam penelitian ini. Sedangkan Observasi ialah menunjukkan sesuatu dengan menggunakan mata. Observasi atau yang disebut dengan pengamatan meliputi kegiatan pemusatan penelitian terhadap suatu objek dengan menggunakan seluruh alat indra. (Dan et al., 2021) Dokumentasi merupakan pengamatan langsung untuk memperoleh data yang sesuai dengan pembahasan. Hal ini dilakukan untuk mengetahui gambaran nyata di lapangan. (Prawiyogi et al., 2021)

Tahapan penelitian

1. Pengumpulan Data, dilakukan dengan cara observasi datang secara langsung ke kantor Desa Cimari
2. Pemilihan Data (Pre-processing), meliputi penghapusan data duplikat, pengecekan data yang tidak konsisten, dan perbaikan kesalahan pada data. Juga dilakukan proses enrichment, artinya proses “memperkaya” data yang ada dengan data atau informasi eksternal.
3. Data transformasi, Pada tahap ini, pemilihan data dilakukan untuk pemilihan atribut. Pada tahap ini ialah melakukan transformasi data yang belum memiliki entitas jelas untuk diubah menjadi bentuk data yang valid atau siap menjalani proses Data Mining.
4. Implementasi K-Means Clustering dengan Microsoft Excel, Dalam melakukan penelitian ini, penulis menggunakan Microsoft excel dan RapidMiner sebagai alat untuk mengolah data, Microsoft Excel digunakan untuk membantu perhitungan. Lalu

yang dilakukan setelah proses perhitungan selesai, dan diterapkan pada RapidMiner untuk memeriksa keakuratan data. Hasil yang diperoleh dari perhitungan di Excel dicocokkan dengan hasil yang diterapkan di RapidMiner.

5. Implementasi K-Means Clustering dengan RapidMiner, adalah salah satu pilihan perangkat lunak untuk mengekstrak data menggunakan data mining. Maka dari itu Pengujian disini menggunakan software Rapidminer

HASIL DAN PEMBAHASAN

3.1 Data Sebelum Preprocessing

Dataset yang diperoleh dari data penduduk lansia Desa Cimari tahun 2021, sebagai berikut :

Tabel 1 Data Penduduk Lansia

Id	JK	Umur	Agama	Alamat
1	Laki-laki	52	Islam	Desa Wetan
2	Laki-laki	64	Islam	Desa Wetan
3	Perempuan	56	Islam	Desa Wetan
.....
1867	Perempuan	51	Islam	Jalatrang
1868	Perempuan	58	Islam	Jalatrang

Memiliki lima (5) atribut : Id, JK, Umur, Agama, dan Alamat. Data penduduk lansia yang diperoleh berjumlah 1.868 record

3.2 Data Transformasi

Dalam proses algoritma K-Means Clustering data yang bisa diolah harus berupa numerik.

Tabel 2 Data Transformasi

Id	JK	Umur	Alamat
1	1	52	Desa Wetan
2	1	64	Desa Wetan
3	2	56	Desa Wetan
.....
299	2	49	Jalatrang
300	2	58	Jalatrang

Memiliki empat (4) atribut yaitu : Id, JK, Umur, dan Alamat. Data yang digunakan sebanyak 300 record.

3.3 Implementasi K-Means Clustering dengan Microsoft Excel

1. Iterasi Awal

Jumlah cluster yang digunakan pada penelitian ini berjumlah 3 cluster yang terdiri dari cluster 1 (C1) lansia awal, cluster 2 (C2) lansia akhir dan cluster 3 (C3) manula. Cluster yang ditentukan menggunakan rumus pengukuran jarak/ *Euclidian Distance* seperti yang terlihat pada tabel dibawah ini.

Tabel 3 Centroid Awal (Secara Random)

Centroid	x	y
C0	2	49
C1	2	56
C2	1	63

Menghitung jarak setiap data terhadap cluster/centroid.

Tabel 4 Hasil Perhitungan Jarak Cluster Iterasi 1

Id	JK	Umur	Alamat	C0	C1	C2	Jarak Terdekat	Cluster
1	1	52	Desa Wetan	3.16	4.12	11	3.16	0
2	1	64	Desa Wetan	15.03	8.06	1	1	2
3	2	56	Desa Wetan	7	0	7.07	0	1
...
299	1	49	Jalatrang	0	7	14.04	0	0
300	1	58	Jalatrang	9	2	5.1	2	1

2. Iterasi Kedua

Tabel 5 Centroid Baru

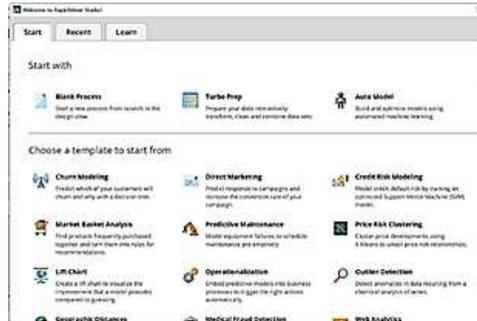
Centroid	x	y
C0	1.546	49.076
C1	1.515	56.079
C2	1.438	63

Tabel 6 Hasil Perhitungan Jarak Cluster Iterasi 2

Id	JK	Umur	Alamat	C0	C1	C2	Jarak Terdekat	Cluster
1	1	52	Desa Wetan	2.97	4.11	11.01	2.97	0
2	1	64	Desa Wetan	14.93	7.94	1.09	1.09	2
3	2	56	Desa Wetan	6.94	0.49	0.49	0.49	1
...
299	1	49	Jalatrang	0.46	7.10	14.01	0.46	0
300	1	58	Jalatrang	8.94	1.98	1.98	1.98	1

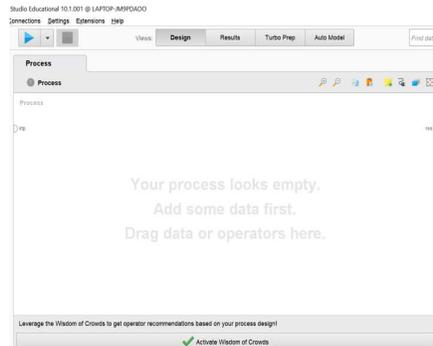
3.4 Implementasi Menggunakan RapidMiner

Pada gambar di bawah merupakan halaman utama dari aplikasi RapidMiner v.10.01 saat pertama kali membuka aplikasi setelah proses loading selesai terlihat pada gambar berikut:



Gambar 2 Tampilan Utama RapidMiner

Kemudia pilih Blank Process untuk membuat halaman kerja pada RapidMiner

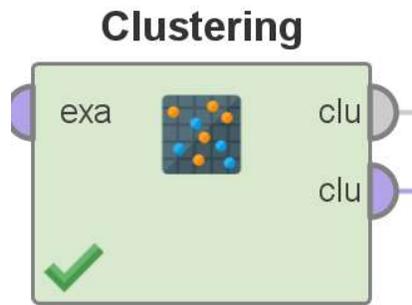


Gambar 3 Tampilan Blank Process

Read Excel

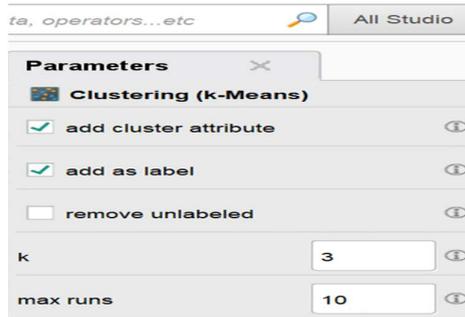


Gambar 4 Operator ReadExcel

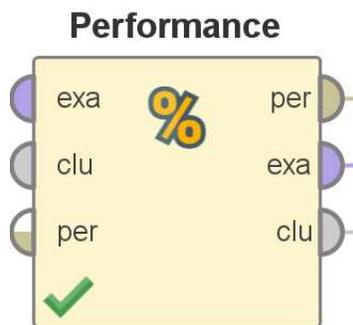


Gambar 5 Operator Clustering

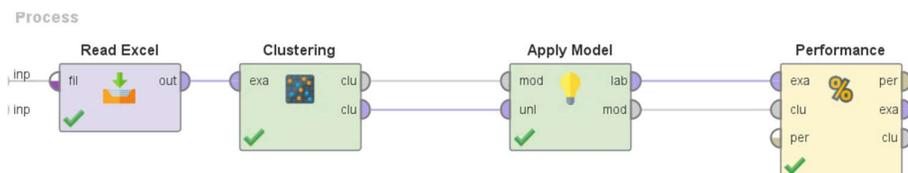
Selanjutnya melakukan pengaturan jumlah cluster, yang diatur pada menu parameters seperti terlihat pada gambar.



Gambar 6 Penentuan Jumlah Cluster



Gambar 7 Operator Performance



Gambar 8 Pemodelan Algoritma K-Means Clustering

Cluster Model (Cluste

Cluster Model

Cluster 0: 119 items
 Cluster 1: 101 items
 Cluster 2: 80 items
 Total number of items: 300

Gambar 9 Cluster Model

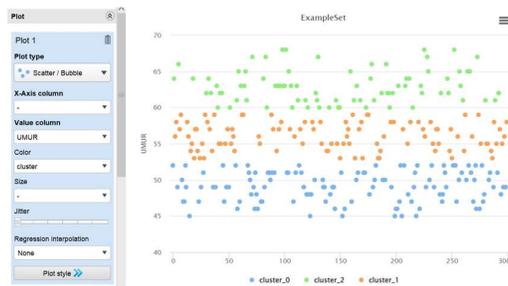
Attribute	cluster_0	cluster_1	cluster_2
JK	1.546	1.515	1.438
UMUR	49.076	56.079	63

Gambar 10 Centroid Table

Row No.	Id	label	ALAMAT	JNSKLMIN	UMUR
1	1	cluster_0	Desa Wetan	1	52
2	2	cluster_2	Desa Wetan	1	64
3	3	cluster_1	Desa Wetan	2	56
4	4	cluster_1	Desa Wetan	1	58
5	5	cluster_0	Desa Wetan	2	49
6	6	cluster_2	Desa Wetan	1	66
7	7	cluster_1	Desa Wetan	2	57
8	8	cluster_1	Desa Wetan	1	59
9	9	cluster_0	Desa Wetan	2	50
10	10	cluster_0	Desa Wetan	1	47
11	11	cluster_0	Desa Wetan	1	47
12	12	cluster_0	Desa Wetan	2	52
13	13	cluster_0	Desa Wetan	2	49

ExampleSet (300 examples, 3 special attributes, 2 regular attributes)

Gambar 11 Example Set



Gambar 12 Tampilan Visualizations



Gambar 13 Davies Bouldin

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan perhitungan dan uji coba menggunakan RapidMiner, peneliti menarik beberapa kesimpulan yaitu:

1. Dari tiga cluster yang telah ditentukan, untuk cluster 0 terdapat 119 item, cluster 1 terdapat 101 item, dan cluster 2 terdapat 80 item.
2. Diperoleh hasil Davies Bouldin sebesar 0.263
3. Tidak ada perbedaan hasil akhir antara perhitungan excel dengan RapidMiner.

Saran

Adapun saran yang dapat dikemukakan yaitu: Untuk peneliti selanjutnya dalam mengelompokkan jumlah penduduk sebaiknya menggunakan metode yang lain selain algoritma K-Means Clustering.

DAFTAR REFERENSI

- Andini, T. D., & Farokhah, L. (2022). Peningkatan Ketersediaan Darah Sesuai Segmentasi Umur Menggunakan K-Means Clustering. *Jurnal Manajemen Informatika (JAMIKA)*, 12(2), 126–136. <https://doi.org/10.34010/jamika.v12i2.7897>
- Dan, T., Yudianto, A., & Artikel, I. (2021). PENGARUH MOTIVASI BELAJAR TERHADAP HASIL BELAJAR SISWA KELAS VII SMP NEGERI 2 KEDOKAN BUNDER KABUPATEN INDRAMAYU. *Jurnal Pendidikan Indonesia*, 2(1).
- Dian, O., Putri, E., Keperawatan, J. I., Dharmas, U., Lintas, I. J., Km, S., Koto, K., Kabupaten, B., Propinsi, D., & Barat, S. (2021). HUBUNGAN FUNGSI KOGNITIF DENGAN KUALITAS HIDUP LANSIA. 2(4). <http://undhari.ac.id>
- Homepage, J., Syaputri, D., Herwina Noprita, P., & Romelah, S. (2021). MALCOM: Indonesian Journal of Machine Learning and Computer Science Implemnetation of K-Means Algorithm for Economic Distribution Clustering Base on Demographics of Population Implementasi Algoritma K-Means untuk Pengelompokan Distribusi Sosial Ekonomi Masyarakat Berdasarkan Demografi Kependudukan. 1, 1–6.
- Jurnal, H., Ali, I., Rinaldi Dikananda, A., Ali Ma, F., & Abdurohman, M. (2021). JURNAL MANAJEMEN INFORMATIKA PENGELOMPOKAN JUMLAH PENDUDUK BERDASARKAN KATEGORI USIA 0-18 TAHUN DENGAN MENGGUNAKAN ALGORITMA K-MEANS UNTUK MENENTUKAN PENGEMBANGAN POTENSI DESA WISATA DI KABUPATEN CIREBON. *JUMIKA*, 8(1).
- Prawiyogi, A. G., Sadiyah, T. L., Purwanugraha, A., & Elisa, P. N. (2021). Penggunaan Media Big Book untuk Menumbuhkan Minat Membaca di Sekolah Dasar. *Jurnal Basicedu*, 5(1), 446–452. <https://doi.org/10.31004/basicedu.v5i1.787>
- Ratna Sari, Y., Sudewa, A., Ayu Lestari, D., & Ika Jaya, T. (2020). PENERAPAN ALGORITMA K-MEANS UNTUK CLUSTERING DATA KEMISKINAN PROVINSI BANTEN MENGGUNAKAN RAPIDMINER (Vol. 5, Issue 2).
- Sonang, S., Purba, A. T., & Pardede, F. O. I. (2019). PENGELOMPOKAN JUMLAH PENDUDUK BERDASARKAN KATEGORI USIA DENGAN METODE K-MEANS. *Jurnal Teknik Informasi Dan Komputer (Tekinkom)*, 2(2), 166. <https://doi.org/10.37600/tekinkom.v2i2.115>
- Yulianti, Y., Utami, D. Y., Hikmah, N., & Hasan, F. N. (2019). PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS UNTUK MENGETAHUI MINAT CUSTOMER DI TOKO HIJAB. *Jurnal Pilar Nusa Mandiri*, 15(2), 241–246. <https://doi.org/10.33480/pilar.v15i2.650>