JITET (Jurnal Informatika dan Teknik Elektro Terapan)
Vol. 13 No. 1, pISSN: 2303-0577 eISSN: 2830-7062 http://dx.doi.org/10.23960/jitet.v13i1.5639

PENERAPAN ALGORITMA K-MEANS DALAM ANALISIS DATA KEPENDUDUKAN UNTUK OPTIMALISASI PENGELOMPOKAN DI DESA PASAWAHAN

Rizki Ramdani[1*], Nana Suarna[2], Irfan Ali[3], Dendy Indriya Efendi[4]

[1,2,3,4] STMIK IKMI Cirebon; Jl. Perjuangan No. 10B, Karyamulya, Kesambi, Kota Cirebon, Jawa Barat, Telp: 0231-490480

Received: 9 Desember 2024
Accepted: 14 Januari 2025
Published: 20 Januari 2025

Keywords:
K-Means, Data Mining, DBI, Population Data

Corespondent Email:
riskiram122@gmail.com

Abstrak. Teknologi informasi, bisnis, pendidikan, kesehatan, dan tata kelola pemerintahan adalah beberapa aspek kehidupan yang sangat dipengaruhi oleh kemajuan pesat di bidang informatika. Kurangnya informasi tentang karakteristik masyarakat Desa Pasawahan adalah masalah utama dalam penelitian ini. Kurangnya informasi ini berdampak pada kemampuan untuk membuat keputusan dan mengatur program pembangunan yang lebih sesuai dengan tujuan. Metode Elbow dan evaluasi Davies-Bouldin Index (DBI) adalah contoh metode klasterisasi yang efektif yang berfokus pada aspek sosial dan ekonomi. Data penduduk Desa Pasawahan Ciamis tahun 2023 dengan 4939 data dan 19 atribut digunakan. Metode Knowledge Discovery in Database (KDD) digunakan untuk menganalisis data. Dalam penelitian ini, indeks Davies-Bouldin (DBI) digunakan untuk mengoptimalkan algoritma K-means untuk menentukan jumlah klaster yang ideal. Hasil pengelompokan menunjukkan bahwa dari jumlah klaster dengan nilai K = 2, 3, 4, 5, 6, 7, 8, 9, 10, klaster dengan kinerja terbaik mendekati nilai 0, dengan nilai DBI sebesar 0,3066. Pada K=3, tiga klaster dengan distribusi sebagai berikut dihasilkan: C0 memiliki 2242 data, C1 memiliki 1778 data, dan C2 memiliki 919 data. Evaluasi kinerja Menurut algoritma K-means, nilai Davies-Bouldin yang mendekati nol menunjukkan kualitas, jumlah klaster yang ideal untuk penelitian ini adalah 3, yang memberikan hasil terbaik.

Abstract. *Information technology, business, education, health, and governance are among the aspects of life significantly influenced by rapid advancements in the field of informatics. One major issue addressed in this study is the lack of information regarding the characteristics of the residents of Pasawahan Village. This information gap impacts decision-making and the ability to implement more targeted development programs. The Elbow method and Davies-Bouldin Index (DBI) evaluation are examples of effective clustering methods focusing on social and economic aspects. This study utilized the 2023 population dataset of Pasawahan Village, Ciamis, consisting of 4,939 records and 19 attributes. The Knowledge Discovery in Database (KDD) methodology was applied to analyze the data. The Davies-Bouldin Index (DBI) was employed to optimize the K-means algorithm and determine the ideal number of clusters. The clustering results revealed that, among the tested cluster numbers (K = 2, 3, 4, 5, 6, 7, 8, 9, 10), the best performance, with a DBI value of 0.3066, was achieved at K=3. This resulted in three clusters with the following distribution: C0 contained 2,242 records, C1 contained 1,778 records, and C2 contained 919 records. The evaluation confirmed that a lower DBI value indicates better clustering quality.*

Therefore, the ideal number of clusters for this study is 3, yielding optimal results.

1. PENDAHULUAN

Teknologi informasi, bisnis, pendidikan, kesehatan, dan tata kelola pemerintahan adalah beberapa aspek kehidupan yang sangat dipengaruhi oleh kemajuan informatika yang pesat[1], [2]. Di Indonesia, data science dan pembelajaran mesin semakin digunakan dalam pengambilan keputusan berbasis data, seperti dalam pemetaan potensi dan kebutuhan masyarakat desa. Menurut [3] , penggunaan data di tingkat desa diharapkan akan membantu menemukan pola dan karakteristik masyarakat untuk memaksimalkan alokasi sumber daya dan pelaksanaan program pembangunan desa.

Kurangnya informasi tentang karakteristik masyarakat Desa Pasawahan adalah masalah utama dalam penelitian ini. Membuat keputusan dan pengorganisasian program pembangunan dipengaruhi oleh kekurangan informasi ini. agar data dapat digunakan untuk membuat kebijakan desa yang menggambarkan berbagai aspek kehidupan masyarakat. Desa Pasawahan membutuhkan gambaran mendalam tentang keadaan masyarakatnya.

Dalam beberapa studi sebelumnya, metode K-Means telah digunakan untuk mengelompokkan data berdasarkan karakteristik sosial ekonomi dan demografi, tetapi masing-masing memiliki ruang lingkup yang beragam dan beberapa keterbatasan, yang membuka peluang untuk penelitian lebih lanjut. melakukan analisis klaster pada data pendidikan di Kabupaten Karawang untuk memahami faktor pendukung pendidikan. Namun, analisis tersebut hanya berfokus pada pendidikan dan tidak mempertimbangkan faktor lain seperti status sosial atau pekerjaan[4]. Menggunakan K-Means untuk mengkategorikan penerima bantuan sosial yang terdaftar dalam Program Keluarga Harapan (PKH), tetapi penelitian ini hanya membahas penerima bantuan dan tidak membahas aspek pekerjaan atau faktor lain yang mungkin memengaruhi kebutuhan bantuan mereka[2]. Menekankan distribusi sosial ekonomi berdasarkan demografi menunjukkan bahwa metode K-Means efektif dalam menentukan kelompok masyarakat yang membutuhkan perhatian khusus[5]. Memperluas penggunaan K-Means untuk mengelompokkan masyarakat berdasarkan berbagai karakteristik sosial yang lebih luas dapat menghasilkan hasil analisis yang lebih kaya dan beragam.

Tujuan penelitian ini adalah untuk membuat model klasterisasi yang bergantung pada algoritma K-means untuk mengelompokkan masyarakat Desa Pasawahan. Diharapkan bahwa dengan membuat klaster yang lebih luas, penelitian ini akan berfungsi sebagai dasar untuk perencanaan pembangunan desa berbasis data yang lebih tepat sasaran.

Algoritma K-means adalah dasar dari penelitian ini, yang digunakan untuk menggabungkandata penduduk Desa Pasawahan ke dalam kelompok untuk analisis sosial-ekonomi. Setelah mengumpulkan data demografis dan sosial-ekonomi, algoritma ini digunakan untuk mengelompokkan data untuk analisis sosial-ekonomi. Jumlah klaster yang ideal dapat dihitung dengan menggunakan metode Elbow dan evaluasi Davies-Bouldin Index (DBI).

2. TINJAUAN PUSTAKA

2.1 Data Penduduk

Untuk mendorong pembangunan sosial dan ekonomi di suatu daerah, pengelolaan data penduduk yang efektif dan akurat sangat penting. Studi menunjukkan bahwa teknologi data dapat meningkatkan akurasi dan kecepatan pengolahan data demografi. Penelitian Nurwati menunjukkan salah satu contoh penerapan ini, menekankan bahwa registrasi penduduk yang sah dan dapat diandalkan sangat penting untuk perencanaan pembangunan[6].

2.2 Clustering

Data mining adalah suatu proses yang mengekstrak informasi berguna dari sejumlah besar data dengan menggunakan berbagai teknik dan algoritma untuk menemukan pola atau pengetahuan yang tersembunyi dalam data, yang membantu membuat keputusan yang lebih baik dan mendapatkan pemahaman yang lebih baik tentang apa yang ada di dalam data[7]. Untuk mengelompokkan kumpulan data

menjadi sejumlah kelompok yang terdiri dari entitas atau objek yang memiliki karakteristik yang serupa, teknik data mining yang dikenal sebagai clustering digunakan[8].

2.3 K-Means

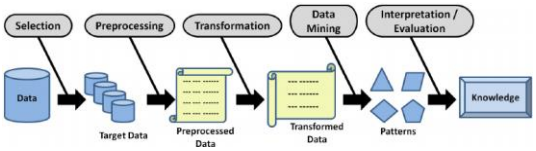
Teknik pengelompokan data non-hirarkisasi digunakan untuk memecah data ke dalam satu atau lebih kluster atau kelompok dengan algoritma K-means[9]. Metode clustering yang dievaluasi dengan K-Means dapat memberikan hasil yang optimal dan akurat untuk menyatukan data yang memiliki ciri yang serupa dan berbeda dari yang lain[10].

2.4 DBI

Penelitian menggunakan teknik kinerja Davies Bouldin Index [11].Nilai DBI yang rendah, atau sama sekali tidak ada, menunjukkan kualitas kluster yang dihasilkan [12].

3. METODE PENELITIAN

Knowledge Knowledge Discovery in Databases (KDD) adalah teknik menemukan pengetahuan dalam dataset. KDD adalah proses mengekstraksi atau mengidentifikasi pola, pengetahuan, dan informasi dari kumpulan data yang besar. Proses KDD inovatif, mudah dipahami, dan bermanfaat bagi pengguna data[13].



Gambar 1. Tahapan KDD

3.1 Data Selection

Data demografis Desa Pasawahan Data sekunder yang digunakan dalam penelitian ini berasal dari kantor Desa Pasawahan, yang memiliki 4939 baris dan 19 atribut.

3.2 Preprocessing Data

Preprocessing data, juga dikenal sebagai data cleaning, adalah proses membersihkan data dari elemen yang tidak penting seperti data duplikat, nilai yang hilang, dan outlier. Proses ini dilakukan dengan mengubah tipe data dari sumber data sebelumnya untuk menghapus data duplikat, nilai yang hilang, dan typo.

3.3 Data Transformation

Transformasi data adalah proses mengubah jenis data sehingga sesuai dengan analisis yang akan dilakukan. Pada atribut status, data dalam penelitian ini diubah dari non-numerik menjadi numerik.

3.4 Data Mining

Tahap awal proses data mining, yang merupakan inti dari proses KDD, dilakukan dengan algoritma clustering K-means. Selain itu, untuk melakukan evaluasi kinerja, Mengenai berapa banyak cluster yang ideal, Davies Bouldin Index (DBI) digunakan. Selanjutnya, proses iteratif dilakukan untuk menghasilkan nilai keanggotaan yang stabil untuk setiap kelompok data mining.

3.5 Evaluation

Merupakan proses melihat atau menilai pola yang sudah ada sebelumnya. Ini dilakukan untuk menentukan apakah pola memberikan hasil yang diharapkan dan membantu dalam pengambilan keputusan.

3.6 Knowledge

Tahap ini melibatkan analisis menyeluruh terhadap pola, tren, atau hubungan dalam data. Tujuan dari tahap ini adalah untuk membuat kesimpulan yang relevan dan membuat keputusan yang didukung oleh informasi yang diperoleh dari analisis yang dilakukan pada tahap sebelumnya.

.

4. HASIL DAN PEMBAHASAN

Dalam penelitian ini, metodologi Knowledge Discovery in Databases (KDD) digunakan.

4. 1 Data Selection

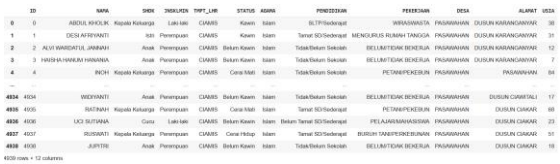
Data demografis Desa Pasawahan digunakan dalam penelitian ini. Data sekunder ini diperoleh dari kantor Desa Pasawahan dan terdiri dari 4939 record dengan 19 atribut. Hanya 12 atribut yang digunakan, yaitu identitas, nama, shdk, jenis kelamin, tempat lahir, status, agama, pendidikan, pekerjaan, desa, alamat, usia, dan kategori usia, karena

pilihan dataset yang tidak digunakan hanya menggunakan 7 atribut.

```
Data_selection=data.drop(columns=['GDR','TGL_LHR',
                                  'KODE_POS','TAHUN_LAHIR',
                                  'TAHUN_SEKARANG','NO_RT','NO_RW'])
Data_selection
```

Gambar 2. Drop Atribut

Dalam data selection terdapat 7 atribut yang dihapus karena tidak memiliki korelasi yang bagus untuk penelitian yang sedang dilakukan



Gambar 3. Hasil Drop Atribut

4.2 Preprocessing

untuk memastikan bahwa data akurat dan relevan dengan tujuan penelitian, berbagai prosedur dilakukan selama tahap pre-processing data. Berikut ini adalah penjelasan tentang proses pemrosesan data sebelum menggunakannya. Data akan diperiksa untuk nilai yang salah, typo, atau duplikat setelah proses seleksi data.

4.3 Transformation

Proses transformasi data dimulai setelah tahapan preprocessing data selesai. Pada tahap transformasi, atribut baru ditambahkan untuk mempermudah pengelompokan.

```
from sklearn.preprocessing import LabelEncoder
le=LabelEncoder()
Data_selection['le_JENIS_KELAMIN']=le.fit_transform(Data_selection['JENIS_KELAMIN'])
Data_selection['le_STATUS']=le.fit_transform(Data_selection['STATUS'])
Data_selection['le_PEKERJAAN']=le.fit_transform(Data_selection['PEKERJAAN'])
Data_selection['le_DESA']=le.fit_transform(Data_selection['DESA'])
Data_selection['le_STATUS_HUBUNGAN_KELUARGA']=le.fit_transform(Data_selection['STATUS_HUBUNGAN_KELUARGA'])
Data_selection['le_ALAMAT']=le.fit_transform(Data_selection['ALAMAT'])
Data_selection['le_Kategori_Usia']=le.fit_transform(Data_selection['Kategori_Usia'])
Data_selection=Data_selection.replace({'PENDIDIKAN': {
                                    'Tidak/Belum Sekolah':0,
                                    'Tamat SD/Sederajat':2,
                                    'Belum Tamat SD/Sederajat':1,
                                    'SLTP/Sederajat':3,
                                    'SLTA/Sederajat':4,
                                    'Diploma I/II':5,
                                    'Diploma III/S. Muda':6,
                                   'Diploma IV/Strata I':7,
                                    'Strata II':8,'Strata III':9,
                                    'Akademi/Diploma III/S. Muda':10
}})
nilai=Data_selection[['PENDIDIKAN','le_PEKERJAAN','le_STATUS','le_STATUS_HUBUNGAN_KELUARGA']]
```

Gambar 4. Proses Transformasi Data

Semua data yang bertipe data kategorikal diubah menjadi nilai numerik dengan tujuan untuk melihat korelasi setiap atribut agar proses pengelompokan jauh lebih optimal.

	PENDIDIKAN	le_PEKERJAAN	le_STATUS	le_STATUS_HUBUNGAN_KELUARGA
0	3	24	3	2
1	2	11	3	1
2	0	0	0	0
3	0	0	0	0
4	0	20	2	2
...	...	...	...	...
4934	0	0	0	0
4935	2	20	2	2
4936	1	15	0	3
4937	2	20	1	2
4938	0	0	0	0

4939 rows × 4 columns

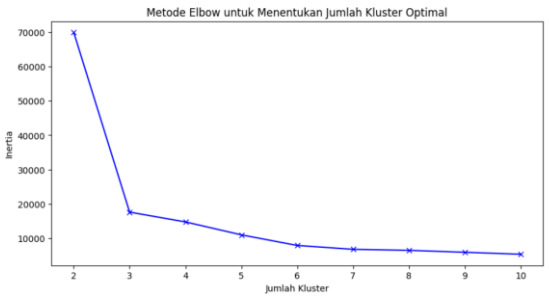
Gambar 5. Hasil Transformasi Data

4. 4 Data Mining

Langkah selanjutnya adalah menghitung jumlah cluster yang ada. Nilai K digunakan untuk menentukan berapa banyak anggota dan klaster yang ideal untuk setiap klaster. Menggunakan metode elbow. Metode Elbow diberi nama karena kurva plot yang dihasilkannya menunjukkan lengkungan yang menyerupai siku pada bagian siku. Jumlah klusteer yang ideal membentuk lengkungan yang menyerupai siku pada grafik plot.

```
[334] import seaborn as sns
      import matplotlib.pyplot as plt
      import plotly.express as px
      from sklearn.cluster import KMeans
      from sklearn.metrics import silhouette_score
      from sklearn.metrics import davies_bouldin_score
      inertia = []
      silhouette_scores = []
      K = range(2, 11)  # Kluster minimal 2
      for k in K:
          kmeans = KMeans(n_clusters=k, random_state=42)
          kmeans.fit(nilai)
          inertia.append(kmeans.inertia_)
          labels = kmeans.labels_
          silhouette_avg = silhouette_score(nilai, labels)
          dbi = davies_bouldin_score(nilai, labels)
          silhouette_scores.append(silhouette_avg)
      plt.figure(figsize=(10, 5))
      plt.plot(K, inertia, 'bx-')
      plt.xlabel('Jumlah Kluster')
      plt.ylabel('Inertia')
      plt.title('Metode Elbow untuk Menentukan Jumlah Kluster Optimal')
      plt.show()
```

Gambar 6. Mencari K Berdasarkan Metode Elbow



Gambar 7. Hasil Proses Metode Elbow

Gambar 7 menunjukkan bahwa titik siku terletak di cluster 2. yang menunjukkan bahwa jumlah cluster yang ideal adalah dua cluster.

Pada tahap berikutnya, proses clustering K-Means dimulai dengan menghitung nilai centroid untuk masing-masing klaster menggunakan data yang tersedia.

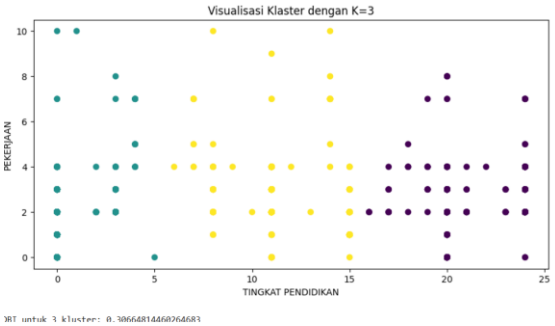
4.5 Interpretation/Evaluation

Pada tahap evaluasi ini, kluster yang telah dibangun akan diperiksa kualitasnya menggunakan pendekatan DBI ini untuk menentukan kekuatan atau kualitasnya.

```
337] kmeans = KMeans(n_clusters=3, random_state=0)
     Data_selection['Cluster'] = kmeans.fit_predict(nilai)
     silhouette_avg = silhouette_score(nilai, Data_selection['Cluster'])
     dbi = davies_bouldin_score(nilai, Data_selection['Cluster'])
     plt.figure(figsize=(10.3, 5))
     plt.scatter(nilai.iloc[:, 1], nilai.iloc[:, 0], c=Data_selection['Cluster'],
                 cmap='viridis')
     plt.title('Visualisasi Klaster dengan K=3')
     plt.xlabel('TINGKAT PENDIDIKAN')
     plt.ylabel('PEKERJAAN')
     plt.show()
     print(f'DBI untuk 3 kluster: {dbi}')
```

Gambar 8. Proses Menampilkan K Terbaik

Dalam perhitungan DBI di atas, kami menemukan nilai 0,3066. Menurut aturan DBI, cluster yang dihasilkan lebih baik jika nilai DBI mendekati 0 atau lebih rendah.



Gambar 9. Hasil K Terbaik

4.6 Knowledge

Hasil analisis menunjukkan bahwa klasterisasi berhasil menghasilkan cluster dengan kualitas terbaik pada K=3, atau 0.3066, dengan nilai DBI paling mendekati 0.

Data dikelompokkan ke dalam kelompok-kelompok berdasarkan kemiripan atribut mereka. Pengelompokan ini dilakukan dengan menggunakan teknik K-Means Clustering. Menurut gambar di atas, berikut adalah deskripsi atribut masing-masing kelompok:

Sebagian besar anggota cluster 0 memiliki karakteristik yang signifikan; ini terutama termasuk gelar SD dan pekerjaan sebagai petani. Sebagian besar dari mereka adalah laki-laki, menikah, dan berfungsi sebagai kepala keluarga. Kelompok lanjut usia mendominasi kelompok ini, rata-rata 55,35 tahun.

Sebagian besar anggota cluster 1 memiliki karakteristik utama: mayoritas belum sekolah, belum atau tidak bekerja, dan sebagian besar adalah laki-laki, belum kawin, dan berfungsi sebagai anak dalam keluarga. Cluster ini mencakup orang dewasa dan remaja, dengan rata-rata usia 19,94 tahun.

Anggota cluster 2 memiliki karakteristik utama, yaitu sebagian besar sedang sekolah dasar dan sudah lulus SD, sebagian besar adalah ibu rumah tangga dengan jenis kelamin perempuan, dan sebagian besar memiliki status sosial kawin. Sebagian besar dari anggota cluster 2 memiliki status hubungan keluarga sebagai istri. Orang-orang dalam kategori usia ini termasuk orang dewasa dan orang lanjut usia, dengan usia rata-rata 41,41 tahun.

Tidak hanya relevansi harus dibahas secara menyeluruh, tetapi pembahasan juga harus menjelaskan bagaimana hasil penelitian dibandingkan dengan temuan yang relevan. Akibatnya, kutipan penting harus dimasukkan ke dalam bagian diskusi. Pada bagian terakhir, hasil penelitian keilmuan harus dibahas dengan

5. KESIMPULAN

a. Data penduduk Desa Pasawahan dianalisis dengan metode clustering K-Means, yang membagi data ke dalam beberapa kelompok berdasarkan seberapa mirip nilainya dengan fitur. Data ini difokuskan pada kategori usia, tingkat pendidikan, pekerjaan, status perkawinan, jenis kelamin, status hubungan keluarga, dan status pendidikan. Pengujian dimulai dengan nilai K mulai dari 2 hingga 10. Performa setiap pembagian cluster dinilai dengan menghitung nilai DBI.

b. Hasil pengujian menunjukkan bahwa nilai DBI yang paling dekat dengan nol menunjukkan kualitas klasterisasi yang tinggi. Nilai DBI tertinggi diperoleh pada saat K=3 dengan nilai DBI 0,3066, yang menunjukkan bahwa nilai optimal untuk pembagian cluster pada data penduduk Desa Pasawahan adalah K=3, yang menghasilkan cluster yang lebih kompak dan terpisah dari cluster lainnya.

c. Hasil dari pengelompokan memberikan gambaran tentang komposisi penduduk

berdasarkan tingkat pendidikan, pekerjaan, status pernikahan, jenis kelamin, status hubungan keluarga, dan kategori usia yang dominan di tiap kelompok. Dengan mengetahui karakteristik masing-masing kelompok, perencanaan kebijakan dapat dilakukan dengan lebih baik dan lebih terarah.

a) Cluster 0 terdiri dari orang-orang yang memiliki gelar SD dan bekerja sebagai petani. Untuk meningkatkan produktivitas sektor pertanian, kelompok ini membutuhkan program pengembangan keterampilan pertanian atau akses ke pelatihan yang relevan.
b) Cluster 1: Banyak orang di sana tidak memiliki pekerjaan atau pendidikan formal. Agar mereka lebih siap untuk menghadapi dunia kerja, kelompok ini dapat menjadi target utama untuk program pendidikan dasar dan pelatihan keterampilan kerja.
c) Cluster 2 terdiri dari individu yang sudah menikah dan bertanggung jawab atas rumah tangga mereka. Untuk meningkatkan kesehatan dan peluang ekonomi keluarga, program pendidikan lanjutan atau pelatihan keterampilan kewirausahaan adalah pilihan yang tepat.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang berpartisipasi dalam penelitian ini.

DAFTAR PUSTAKA

[1] N. A. Maori and E. Evanita, "Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 2, pp. 277–288, 2023, doi: 10.24176/simet.v14i2.9630.

[2] S. Abdul, Azis, S. Defit, and Y. Yunus, "Klasterisasi Dana Bantuan Pada Program Keluarga Harapan (PKH) Menggunakan Metode K-Means," *J. Inform. Ekon. Bisnis*, vol. 3, no. 2, pp. 53–59, 2021, doi: 10.37034/infeb.v3i2.66.

[3] M. Darwis, L. H. Hasibuan, M. Firmansyah, N. Ahady, and R. Tiaharyadini, "Implementation of K-Means clustering algorithm in mapping the groups of graduated or dropped-out students in the Management Department of the National University," *JISA(Jurnal Inform. dan Sains)*, vol. 4, no. 1, pp. 1–9, 2021, doi: 10.31326/jisa.v4i1.848.

[4] Abdussalam Amrullah, Intam Purnamasari, Betha Nurina Sari, Garno, and Apriade Voutama, "ANALISIS CLUSTER FAKTOR PENUNJANG PENDIDIKAN MENGGUNAKAN ALGORITMA K-MEANS (STUDI KASUS: KABUPATEN KARAWANG)," *J. Inform. dan Rekayasa Elektron.*, vol. 5, no. 2, pp. 244–252, 2022, doi: 10.36595/jire.v5i2.701.

[5] D. Syaputri, P. H. Noprita, and S. Romelah, "Implementasi Algoritma K-Means untuk Pengelompokan Distribusi Sosial Ekonomi Masyarakat Berdasarkan Demografi Kependudukan," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 1–6, 2021, doi: 10.57152/malcom.v1i1.5.

[6] E. A. Herdiaman, A. Sudiarjo, M. Hikmatyar, T. Informatika, U. Perjuangan, and J. Barat, "RSUD MENGGUNAKAN METODE K-MEANS," vol. 12, no. 3, 2024.

[7] R. R. Rerung, "Penerapan Data Mining dengan Memanfaatkan Metode Association Rule untuk Promosi Produk," *J. Teknol. Rekayasa*, vol. 3, no. 1, p. 89, 2018, doi: 10.31544/jtera.v3.i1.2018.89-98.

[8] E. Muningsih, I. Maryani, and V. R. Handayani, "Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa," *J. Sains dan Manaj.*, vol. 9, no. 1, p. 96, 2021, [Online]. Available: www.bps.go.id

[9] F. Khoirunnisa and Y. Rahmawati, "Komparasi 2 Metode Cluster Dalam Pengelompokan Intensitas Bencana Alam Di Indonesia," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 1, pp. 68–79, 2024, doi: 10.23960/jitet.v12i1.3619.

[10] Ni Putu Viona Viandari, I Made Agus Dwi Suarjaya, and I Nyoman Piarsa, "Pemetaan Pelanggan dengan LRFM dan Two Stage Clustering untuk Memenuhi Strategi Pengelolaan," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 130–139,

2022, doi: 10.29207/resti.v6i1.3778.
[11] M. Sholeh and K. Aeni, "Perbandingan Evaluasi Metode Davies Bouldin, Elbow dan Silhouette pada Model Clustering dengan Menggunakan Algoritma K-Means," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 8, no. 1, p. 56, 2023, doi: 10.30998/string.v8i1.16388.
[12] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, p. 100, 2021.s
[13] A. Yudhistira and R. Andika, "Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering," *J. Artif. Intell. Technol. Inf.*, vol. 1, no. 1, pp. 20–28, 2023, doi: 10.58602/jaiti.v1i1.22.