

Data Mining

Implementasi Algoritma K-Means pada Kasus Kekerasan Anak dan Perempuan Berdasarkan Usia

Rahma Fauziah, Ade Irma Purnamasari

Program Studi Manajemen Informatika, STMIK IKMI Cirebon, Cirebon, Indonesia



INFORMASI ARTIKEL

Diterima Redaksi: 18 Februari 2023
Revisi Akhir: 13 Maret 2023
Diterbitkan Online: 15 Maret 2023

KATA KUNCI

Algoritma K-Means; Data Mining; Kekerasan

KORESPONDENSI

Phone: +62 85860269985
E-mail: fauziahrahma053@gmail.com

A B S T R A K

Kekerasan merupakan sebuah tindakan atau ancaman yang memang sengaja dilakukan oleh individu atau kelompok yang dapat menimbulkan dampak negatif, kerugian psikis, trauma, dan perampasan hak. Kekerasan bisa dalam bentuk fisik atau bisa juga dalam bentuk psikis, kekerasan yang terjadi sering kali menimpa perempuan dan anak. Berdasarkan data Komisi Nasional Anti Kekerasan terhadap Perempuan (Komnas Perempuan), terdapat 299.911 kasus kekerasan terhadap perempuan pada tahun 2020. Tujuan dari penelitian ini untuk mengetahui hasil cluster atau pengelompokan dari kasus kekerasan anak dan perempuan berdasarkan kelompok usia agar dapat memudahkan pemerintah untuk mengetahui rentang usia yang paling sering menjadi sasaran kekerasan. Dataset terkait topik penelitian ini dihasilkan oleh Dinas Pemberdayaan Perempuan Anak dan Keluarga (DP3AKB) yang dikeluarkan dalam periode satu tahun sekali. Data yang digunakan adalah data jumlah korban kasus kekerasan berdasarkan kelompok usia. Dataset terdiri dari 945 data. Penelitian ini menggunakan metode algoritma k-means dengan metode *Knowledge Discovery in Database* dan penentuan cluster menggunakan *Davies Bouldin Index*. Hasil dari penelitian ini adalah membagi ke dalam 6 cluster yaitu : Cluster 0 (C0) yaitu 55 anggota yang paling mendominasi kelompok usia 13-17 tahun berjumlah 21, cluster 1 (C1) yaitu 193 anggota yang paling mendominasi kelompok usi 0-5 tahun berjumlah 70, cluster 2 (C2) yaitu 184 anggota yang paling mendominasi kelompok usia 45-59 tahun berjumlah 64, cluster 3 (C3) yaitu 197 anggota yang paling mendominasi kelompok usia 45-59 tahun berjumlah 70, cluster 4 (C4) yaitu 4 anggota kelompok usia 25-44 tahun, cluster 5 (C5) yaitu 177 anggota yang paling mendominasi kelompok usia 0-5 tahun berjumlah 65.

PENDAHULUAN

Tindakan kekerasan adalah perampasan hak seseorang yang menimbulkan dampak negatif terhadap fisik atau psikis baik disengaja atau tidak dalam bentuk perbuatan lainnya [1]. Kekerasan terjadi karena adanya ketimpangan dalam deferensiasi sosial yang akhirnya menyebabkan terjadi tindak kekerasan. Kekerasan dalam arti luas merujuk pada tindakan fisik atau tindakan psikis yang dilakukan oleh seseorang atau sekelompok individu, baik yang dilakukan secara sengaja, langsung maupun tidak langsung, personal atau struktural[2].

Perempuan dan anak lebih rentan menjadi korban kekerasan, karena sering kali dianggap sebagai sosok yang lemah, dan tidak memiliki kekuatan untuk melawan. Kekerasan bisa berupa kekerasan fisik dan psikis yang bisa berakibat pada kerusakan kondisi fisik dan gangguan pada kondisi mental seseorang. Kasus kekerasan terhadap anak dan perempuan merupakan masalah yang sangat mengkhawatirkan bagi masyarakat, dengan meningkatnya kasus kekerasan yang menimpa anak dan perempuan serta sedikitnya korban yang melaporkan kasus yang dialaminya. Ditambah lagi tipisnya jaminan perlindungan terhadap korban, yang dalam hal ini kaum wanita, yang akhirnya korban lebih banyak bersikap diam, karena mana yang benar dan mana yang salah tidak lagi ada pemahamannya yang serius [3].



Gambar 1. Jumlah perempuan korban kekerasan

Catatan Tahunan (CATAHU) 2021 Komnas Perempuan mengeluarkan data kekerasan terhadap perempuan sepanjang tahun 2020 sebanyak 299.911 kasus, jumlah tersebut menurun dari tahun sebelumnya dan terdiri dari kasus yang ditangani oleh: Pengadilan Negeri/Pengadilan Agama berjumlah 291.677 kasus. Lembaga layanan mitra komnas perempuan sejumlah 8234 kasus. Unit Pelayanan dan Rujukan (UPR) komnas perempuan sebanyak 2389 kasus. Sementara angka kekerasan terhadap perempuan yang tertinggi berdasarkan provinsi yaitu DKI Jakarta (2461 kasus) dan Jawa Barat (1011 kasus)[4]. Sedangkan berdasarkan rilis data Komisi Perlindungan Anak Indonesia (KPAI) jumlah pengaduan kasus perlindungan anak berjumlah 5952 kasus. Tingginya kasus menunjukkan bahwa masalah kekerasan masih belum tertangani secara merata.

Merujuk pada permasalahan yang telah disebutkan diatas, penelitian ini bertujuan untuk mengelompokkan kasus kekerasan anak dan perempuan berdasarkan kelompok usia menggunakan algoritma k-means dengan metode *Knowledge Discovery in Database* dan penentuan cluster menggunakan *Davies Bouldin Index*. Adapun manfaat dari penelitian ini diharapkan dapat membantu Dinas Pemberdayaan Perempuan dan Perlindungan Anak (DP3A) Jawa Barat dalam menekan angka kasus kekerasan anak dan perempuan serta pengendalian kasus sehingga dibutuhkan pengelompokan kasus berdasarkan kelompok usia agar dapat memudahkan pemerintah untuk mengetahui rentang usia yang paling sering menjadi sasaran kekerasan.

Berdasarkan penelitian sebelumnya yang membahas tentang pengelompokan algoritma k-means dengan Bahasa pemrograman python dan penentuan jumlah cluster menggunakan *method elbow*. Algoritma k-means banyak digunakan dalam pengelompokan karena kesederhanaan dan efisiensinya. Algoritma ini relatif lebih terukur dan efisien untuk pengolahan objek dalam jumlah besar karena memiliki ketelitian yang optimal terhadap ukuran objek. Selain memiliki kelebihan, algoritma k-means juga memiliki kekurangan dalam menganalisa dan menentukan jumlah cluster terbaik dalam pengelompokan data pada suatu dataset [1].

TINJAUAN PUSTAKA

Kekerasan

Tindak kekerasan adalah tindakan fisik yang disengaja atau tidak disengaja atau bentuk kegiatan lain yang dapat berdampak negatif terhadap kesehatan fisik dan mental serta hilangnya hak seseorang[2]. Menurut [4] Kekerasan merupakan penggunaan kekuatan yang besar secara tiba-tiba dalam keadaan emosi yang tidak terkendali yang dapat mengakibatkan trauma, psikis, kematian, dan kelainan bahkan perampasan hak.

Data Mining

Data mining merupakan metode yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk menghasilkan dan mengidentifikasi informasi yang bermanfaat dan sebagai ekstraksi informasi pengetahuan yang terkait berbagai basis data besar[5]. *Data mining* adalah istilah yang dilakukan untuk mendefinisikan penemuan pengetahuan dari database. *Data mining* mengimplementasikan teknik pengenalan pola (teknik matematika dan statistika) untuk mengamati sejumlah besar data yang disimpan dalam memori untuk meendapat informasi [6].

Knowledge Discovery in Database

Knowledge Discovery in Database (KDD) adalah proses pengambilan informasi yang tersembunyi, di mana informasi tersebut sebelumnya tidak dikenal. KDD merupakan metode yang digunakan untuk mendapat pengetahuan yang berasal dari database yang ada. Hasil pengetahuan yang diterima dapat dimanfaatkan untuk basis pengetahuan (*knowledge base*) yang dipergunakan dalam keperluan mengambil keputusan [7]. *Knowledge Discovery in Database* (KDD) adalah suatu teknik pembentukan pola atau rule dalam informasi. Informasi yang dihasilkan didapatkan dari suatu data yang besar atau dikenal dengan tambang data yang disimpan dalam basis data yang awalnya belum diketahui dan menghasilkan suatu data yang potensial bermanfaat [8].

Clustering

Clustering termasuk salah satu metode dari data mining. *Clustering* adalah suatu proses pengelompokan kelas yang mempunyai kesamaan objek, record, atau observasi [9]. *Clustering* merupakan tahapan untuk mengelompokkan himpunan data yang atribut kelasnya belum diprediksikan. Maka *clustering* pada dasarnya akan menghasilkan sejumlah cluster (kelompok) [10]. Sedangkan menurut [11] *clustering* merupakan metode yang digunakan dalam data mining yang cara kerjanya mencari dan mengelompokkan data yang mempunyai kemiripan karakteristik antara data satu dengan data lainnya yang telah diperoleh.

Algoritma K-Means

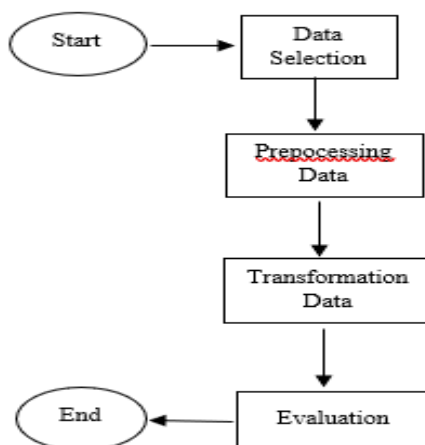
Algoritma k-means adalah algoritma yang membagi data menjadi cluster maka data yang sejenis berada di cluster sama, yang memiliki ketidaksamaan berada di cluster berbeda [6]. Algoritma k-means adalah algoritma non hirarki yang berasal dari metode data *clustering*. Algoritma k-means dimulai dengan membuat bagian cluster diawal kemudian secara berulang cluster ini diperbaiki hingga tidak terjadi Istilah dalam algoritma k-means *clustering* [12].

RapidMiner

RapidMiner adalah perangkat lunak yang bersifat terbuka. RapidMiner merupakan sebuah penyelesaian untuk melakukan analisis tentang data mining, dan analisis prediksi. RapidMiner adalah software yang berdiri sendiri untuk analisis data dan menjadi mesin data mining yang mungkin diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan memakai bahasa java sehingga dapat bekerja di semua sistem operasi [13].

METODOLOGI

Pada tahap ini menjelaskan tentang metode penelitian yang menjelaskan tentang tahapan perancangan yang digunakan. Tahapan perancangan terdiri dari lima tahapan yang dilakukan secara berurutan. Tahapan tersebut terdiri dari *selection data*, *preprocessing data*, *transformation data*, *data mining*, *evaluation*. Adapun penjelasan dari tahap tersebut yaitu:



Gambar 2. Tahap Perancangan

Selection Data

Tahap pertama adalah proses *selection* data yang akan digunakan dalam proses pengelompokan data. Dataset yang digunakan berasal dari data publik melalui situs terbuka opendata.jabarprov.go.id yang dirilis oleh Dinas Pemberdayaan

Perempuan, Perlindungan Anak Dan Keluarga Berencana. Data yang digunakan adalah data jumlah korban kekerasan berdasarkan kelompok usia di Jawa Barat tahun 2017-2021. Dataset terdiri dari 945 data dengan 9 atribut. Atribut tersebut diantaranya id, kode provinsi, nama provinsi, kode kabupaten/kota, nama kabupaten/kota, kelompok usia, jumlah korban, tahun. Adapun hasil dataset setelah dilakukan penghapusan atribut yang tidak diperlukan adalah nama kabupaten/kota, kelompok usia, jumlah korban, dan tahun.

Tabel 1. Dataset Awal

| Id | Kode provinsi | Nama provinsi | Kode kabupaten kota | Nama kabupaten kota | Kelompok usia | Jumlah korban | tahun |
|-----------|----------------------|----------------------|----------------------------|----------------------------|----------------------|----------------------|--------------|
| 1 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | 0 – 5 TAHUN | 0 | 2017 |
| 2 | 32 | JAWA BARAT | 3201 | KABUPATEN BOGOR | 6-12 TAHUN | 1 | 2017 |
| | | | | | | | |
| 943 | 32 | JAWA BARAT | 3279 | KOTA BANJAR | 25-44 TAHUN | 4 | 2021 |
| 944 | 32 | JAWA BARAT | 3279 | KOTA BANJAR | 45-59 TAHUN | 0 | 2021 |
| 945 | 32 | JAWA BARAT | 3279 | KOTA BANJAR | 60+ TAHUN | 0 | 2021 |

Tabel 2. Dataset Hasil Seleksi

| Nama kabupaten kota | Kelompok usia | Jumlah korban | Tahun |
|----------------------------|----------------------|----------------------|--------------|
| KABUPATEN BOGOR | 0-5 TAHUN | 0 | 2017 |
| KABUPATEN BOGOR | 6-12 TAHUN | 0 | 2017 |
| | | | |
| KOTA BANJAR | 13-17 TAHUN | 2 | 2021 |
| KOTA BANJAR | 18-24 TAHUN | 4 | 2021 |
| KOTA BANJAR | 25-44 TAHUN | 0 | 2021 |
| KOTA BANJAR | 45-59 TAHUN | 0 | 2021 |

Preprocessing Data

Tahap selanjutnya *Preprocessing data* yaitu dilakukan penghapusan kelompok umur 60 tahun karena semua jumlah korban bernilai nol serta pengecekan dataset dengan menghilangkan data *missing value*. Pada dataset ini tidak terdapat *missing value* sehingga dapat beralih ke tahap berikutnya yaitu tahap *transformation data*.

| Name | Type | Missing | Statistics | Filter (6 / 6 attributes): |
|---------------------|---------|---------|--|----------------------------|
| id | Integer | 0 | Min: 1, Max: 810, Average: 405.500 | Search for Attributes |
| cluster | Nominal | 0 | Least: cluster_4 (4), Most: cluster_3 (197), Values: cluster_3 (197) | |
| nama_kabupaten_kota | Real | 0 | Min: -1.668, Max: 1.668, Average: -0 | |
| kelompok_usia | Real | 0 | Min: -1.463, Max: 1.463, Average: 0 | |
| jumlah_korban | Real | 0 | Min: -0.599, Max: 8.868, Average: -0.000 | |
| tahun | Real | 0 | Min: 0, Max: 0, Average: 0 | |

Gambar 3. Missing Value

Transformation Data

Tahapan ini berfokus untuk mengubah data yang telah melewati tahapan *preprocessing* agar bisa diproses oleh algoritma dengan mengubah data menjadi bentuk vektor angka [14]. Proses *transformation data* pada penelitian ini yaitu dilakukan inisialisasi dataset atau perubahan nama variabel untuk memudahkan proses pada rapidminer. Penggantian nama variabel yaitu pada kelompok usia diubah menjadi kode angka.

Tabel 3. Inisialisasi Dataset

| KELOMPOK USIA | INISIALISASI |
|---------------|--------------|
| 0-5 TAHUN | 1 |
| 6-12 TAHUN | 2 |
| 13-17 TAHUN | 3 |
| 18-24 TAHUN | 4 |
| 25-44 TAHUN | 5 |
| 45-59 TAHUN | 6 |

Tabel 4. *Transformation Data*

| Nama kabupaten kota | Kelompok usia | Jumlah korban | tahun |
|---------------------|---------------|---------------|-------|
| KABUPATEN BOGOR | 1 | 0 | 2017 |
| KABUPATEN BOGOR | 2 | 0 | 2017 |
| KABUPATEN BOGOR | 3 | 0 | 2017 |
| KABUPATEN BOGOR | 4 | 1 | 2017 |
| KABUPATEN BOGOR | 5 | 1 | 2017 |
| | | | |
| KOTA BANJAR | 1 | 1 | 2021 |
| KOTA BANJAR | 2 | 3 | 2021 |
| KOTA BANJAR | 3 | 11 | 2021 |
| KOTA BANJAR | 4 | 2 | 2021 |
| KOTA BANJAR | 5 | 4 | 2021 |

Data Mining

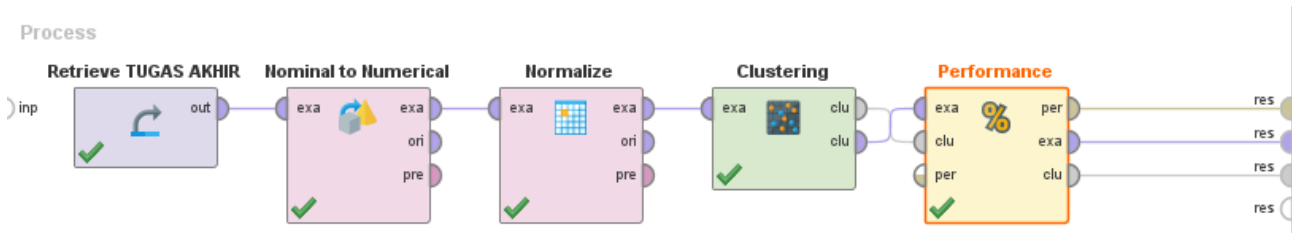
Pada tahap ini merupakan pemilihan metode data mining. Terdapat beberapa metode pada data mining diantaranya yaitu klasifikasi, kalstering asosiasi, regresi, dan lain-lain. Pada penelitian ini metode yang digunakan yaitu kalstering menggunakan algoritma k-means pada tools rapidminer. Algoritma k-means merupakan salah satu algoritma dari *clustering* yang bertujuan untuk mengelompokan data pada kasus kekerasan anak dan perempuan.

Evaluation

Evaluation clustering dilakukan dengan tujuan untuk mengetahui seberapa baik kualitas dari hasil *clustering* [15]. Pada penelitian ini, evaluasi hasil *clustering* yang digunakan adalah *Davies Bouldin Index* untuk mengetahui jumlah cluster yang paling optimal.

HASIL DAN PEMBAHASAN

Berikut tampilan susunan operator yang digunakan dalam pengelompokan kasus kekerasan anak dan perempuan dengan model algoritma k-means.



Gambar 4. Model Tools RapidMiner

Hasil dari implementasi algoritma k-means pada kasus kekerasan anak dan perempuan pada proses pengelompokan tahap awal yaitu menentukan jumlah cluster dengan cara melihat persentase hasil *Davies Bouldin Index* (DBI). Setelah proses *running* pada rapidminer di dapat DBI:

Tabel 5. Nilai *Davies Bouldin Index*

| Cluster | DBI |
|-----------|--------------|
| K2 | 0.365 |
| K3 | 0.281 |
| K4 | 0.234 |
| K5 | 0.225 |
| K6 | 0.211 |
| K7 | 0.225 |
| K8 | 0.245 |
| K9 | 0.250 |
| K10 | 0.259 |

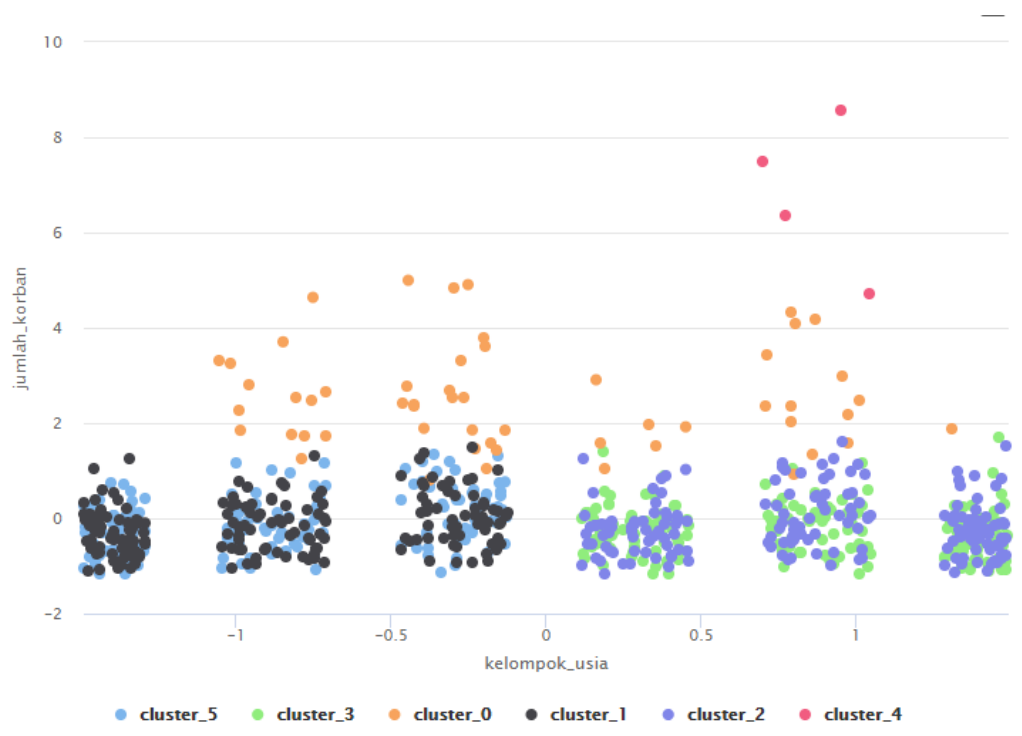
Dari hasil pengolahan diketahui, untuk K optimum berada di cluster 6 dengan DBI 0.211.

Cluster Model

```
Cluster 0: 55 items
Cluster 1: 193 items
Cluster 2: 184 items
Cluster 3: 197 items
Cluster 4: 4 items
Cluster 5: 177 items
Total number of items: 810
```

Gambar 5. Cluster Model

Berdasarkan hasil cluster model diatas dengan DBI 0.211. Cluster yang terbentuk diantaranya cluster 0 (C0) berjumlah 55 anggota, cluster 1 (C1) berjumlah 193 anggota, cluster 2 (C2) berjumlah 184, cluster 3 (C3) berjumlah 197, cluster 4 (C4) berjumlah 4 anggota, Cluster 5 (C5) berjumlah 177 anggota. Hasil dari pengelompokan dapat dilihat pada *scatter* berikut:



Gambar 6. Persebaran Cluster dengan Scatter

Berdasarkan hasil persebaran scatter dan tabel diatas Cluster 0 (C0) yaitu 55 anggota yang paling mendominasi kelompok usia 13-17 tahun berjumlah 21, cluster 1 (C1) yaitu 193 anggota yang paling mendominasi kelompok usi 0-5 tahun berjumlah 70, cluster 2 (C2) yaitu 184 anggota yang paling mendominasi kelompok usia 45-59 tahun berjumlah 64, cluster 3 (C3) yaitu 197 anggota yang paling mendominasi kelompok usia 45-59 tahun berjumlah 70, cluster yaitu 4 anggota kelompok usia 25-44 tahun, cluster 5 (C5) yaitu 177 anggota yang paling mendominasi kelompok usia 0-5 tahun berjumlah 65.

KESIMPULAN DAN SARAN

Berdasarkan hasil pengelompokan kasus kekerasan anak dan perempuan menggunakan algoritma k-means yaitu: Pengelompokan berdasarkan usia terbagi menjadi 6 cluster berdasarkan kemiripan karakteristik kasus kekerasan anak dan perempuan. Adapun keenam cluster tersebut antara lain: Cluster 0 (C0) yaitu 55 anggota yang paling mendominasi kelompok usia 13-17 tahun berjumlah 21, cluster 1 (C1) yaitu 193 anggota, cluster 2 (C2) yaitu 184 anggota yang paling mendominasi kelompok usia 45-59 tahun berjumlah 64, cluster 3 (C3) yaitu 197 anggota yang paling mendominasi kelompok usia 45-59 tahun berjumlah 70, cluster yaitu 4 anggota kelompok usia 25-44 tahun, cluster 5 (C5) yaitu 177 anggota yang paling mendominasi kelompok usia 0-5 tahun berjumlah 65. Hasil dari pengelompokan kasus kekerasan anak dan perempuan dengan jumlah dataset 810 menggunakan algoritma k-means K terbaik yang didapat yaitu berada di cluster 6 dengan DBI 0.211 mendekati 0. Oleh karena itu, untuk pengembangan penelitian selanjutnya, diharapkan agar menggunakan metode atau algoritma lain dalam proses pengelompokan kasus kekerasan anak dan perempuan serta melakukan percobaan dengan parameter yang berbeda untuk mengetahui perbandingan evaluasi yang lebih optimal.

DAFTAR PUSTAKA

- [1] R. Rahma *dkk.*, "PENGELOMPOKAN DAERAH RAWAN KEKERASAN TERHADAP PEREMPUAN DAN ANAK DI JAWA BARAT MENGGUNAKAN ALGORITMA K-MEANS," 2022.
- [2] N. S. M. J. Noviya Adawiyah, "KLAUSTERISASI KASUS KEKERASAN TERHADAP ANAK DAN PEREMPUAN BERDASARKAN ALGORITMA K-MEANS," vol. 5, 2020.
- [3] Andang Sari dan A. Haryani Putri, "Perlindungan Hukum Terhadap Perempuan Korban Kekerasan Dalam Rumah Tangga," *KRTHA BHAYANGKARA*, vol. 14, no. 2, hlm. 236–245, Des 2020, doi: 10.31599/krtha.v14i2.291.

- [4] A. Q. A. Y. Aflina Mustafainah, “data kekerasan seksual pada perempuan (KOMNAS PEREMPUAN),” *KOMNAS PEREMPUAN: PEREMPUAN DALAM HIMPITAN PANDEMI: LONJAKAN KEKERASAN SEKSUAL, KEKERASAN SIBER, PERKAWINAN ANAK, DAN KETERBATASAN PENANGANAN DI TENGAH COVID-19*, 2021.
- [5] B. Septia Pranata dan D. Putro Utomo, “Bulletin of Information Technology (BIT) Penerapan Data Mining Algoritma FP-Growth Untuk Persediaan Sparepart Pada Bengkel Motor (Study Kasus Bengkel Sinar Service),” *Bulletin of Information Technology (BIT)*, vol. 1, no. 2, hlm. 83–91, 2020.
- [6] N. S. M. J. Noviya Adawiyah, “Klasterisasi Kasus Kekerasan Terhadap Anak dan Perempuan Berdasarkan Algoritma K-Means,” 2021.
- [7] G. Gustientiedina, M. H. Adiya, dan Y. Desnelita, “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan,” *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 5, no. 1, hlm. 17–24, Apr 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [8] J. Nasir, “PENERAPAN DATA MINING CLUSTERING DALAM MENGELOMPOKAN BUKU DENGAN METODE K-MEANS,” *Jurnal SIMETRIS*, vol. 11, no. 2, 2020.
- [9] S. Sindi *dkk.*, “ANALISIS ALGORITMA K-MEDOIDS CLUSTERING DALAM PENGELOMPOKAN PENYEBARAN COVID-19 DI INDONESIA,” *Jurnal Teknologi Informasi*, vol. 4, no. 1, 2020.
- [10] R. Muliono dan Z. Sembiring, “DATA MINING CLUSTERING MENGGUNAKAN ALGORITMA K-MEANS UNTUK KLASTERISASI TINGKAT TRIDARMA PENGAJARAN DOSEN,” 2019.
- [11] H. S. Siska Febriani, “ANALISIS DATA HASIL DIAGNOSA UNTUK KLASIFIKASI GANGGUAN KEPERIBADIAN MENGGUNAKAN ALGORITMA C4.5,” vol. 2, hlm. 89–95, 2021.
- [12] A. Sulistiyawati dan E. Supriyanto, “Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan,” vol. 15, no. 2.
- [13] Z. Nabila, A. Rahman Isnain, dan Z. Abidin, “ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS,” *Jurnal Teknologi dan Sistem Informasi (JTISI)*, vol. 2, no. 2, hlm. 100, 2021, [Daring]. Tersedia pada: <http://jim.teknokrat.ac.id/index.php/JTISI>
- [14] A. Habib Husaini, R. Mayasari, dan U. Singaperbangsa Karawang, “PENGELOMPOKAN ULASAN APLIKASI PEDULILINDUNGI DENGAN ALGORITMA K-MEDOIDS PEDULILINDUNGI APPLICATION REVIEW GROUPING WITH THE K-MEDOIDS ALGORITHM,” *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 5, no. 2, 2022.
- [15] E. Muningsih, I. Maryani, dan V. R. Handayani, “Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa,” *Jurnal Sains dan Manajemen*, vol. 9, no. 1, 2021, [Daring]. Tersedia pada: www.bps.go.id