

CLUSTERING KASUS PELECEHAN SEKSUAL ANAK DIBAWAH UMUR DI SULAWESI TENGGARA MENGGUNAKAN K-MEANS

Aditya Putra, Sumiartin, Anya Salsabila, Toga Abdi Hayat, Rafli Balani

Teknik Informatika, Universitas Halu Oleo

Jalan H.E.A. Mokodompit, Kota Kendari, Sulawesi Tenggara

adypptraa@gmail.com

ABSTRAK

Pelecehan seksual terhadap anak di bawah umur merupakan pelanggaran hak asasi manusia yang berdampak serius terhadap perkembangan psikologis dan fisik anak. Maraknya kasus pelecehan seksual di Sulawesi Tenggara menuntut perlunya langkah-langkah penanganan yang lebih efektif. Penelitian ini bertujuan untuk mengelompokkan Kasus kekerasan seksual terhadap anak di bawah umur di Sulawesi Tenggara menggunakan algoritma k-means. Dengan mengelompokkan data kasus, diharapkan dapat ditemukan pola dan karakteristik yang dapat membantu dalam penanganan dan pencegahan kasus serupa di masa depan. Dengan algoritma k-means nilai k optimal yang diperoleh adalah $k=3$ berdasarkan kurva Elbow Method dan mengelompokkan kasus pelecehan seksual anak menjadi tiga cluster yaitu cluster 0 dengan jumlah kabupaten 6 dimana kasus pelecehan seksual cenderung rendah, cluster 1 dengan jumlah kabupaten 3 dimana tingkat kasus pelecehan seksual tinggi, dan cluster 2 dengan jumlah kabupaten 8 dimana tingkat kasus pelecehan seksual sedang. Implementasi penelitian ini berupa aplikasi yang diharapkan dapat digunakan sebagai acuan

Kata kunci : pelecehan, seksual anak, clustering, k-means, Sulawesi Tenggara, aplikasi

1. PENDAHULUAN

Pelecehan seksual terhadap anak di bawah umur adalah salah satu bentuk kejahatan yang paling berbahaya dan traumatis. Pasal 281 hingga 299 KUHP mengatur kejahatan yang berkaitan dengan masalah seksual atau moral. Perilaku seksual yang diungkapkan secara verbal, nonverbal, atau visual disebut pelecehan seksual. Televisi dan media sosial penuh dengan pemberitaan kejahatan seksual terhadap anak yang terjadi hampir setiap hari. Pelecehan seksual terhadap anak-anak, baik perempuan maupun laki-laki, tidak boleh terjadi dalam keadaan apa pun.

Penelitian ini mengusulkan metode K-means untuk menganalisis data terkait pelecehan seksual terhadap anak. K-Means merupakan algoritma clustering yang digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik. Metode ini dipilih untuk memudahkan analisis pola kejadian pelecehan seksual berdasarkan faktor-faktor seperti lokasi, waktu kejadian, usia korban, dan jenis kelamin, sehingga memungkinkan pengelompokan data yang efisien.

Penggunaan K-means diharapkan dapat membantu mengidentifikasi pola-pola tertentu yang umum terjadi pada pelecehan seksual terhadap anak. Hasil analisis ini dapat memberikan wawasan yang lebih mendalam kepada pihak berwenang untuk mengembangkan strategi pencegahan dan tindakan intervensi yang lebih efektif dan berbasis data. Selain itu, hasil pengelompokan ini dapat digunakan sebagai dasar kampanye pendidikan dan kesadaran untuk mencegah pelecehan seksual terhadap anak.

2. TINJAUAN PUSTAKA

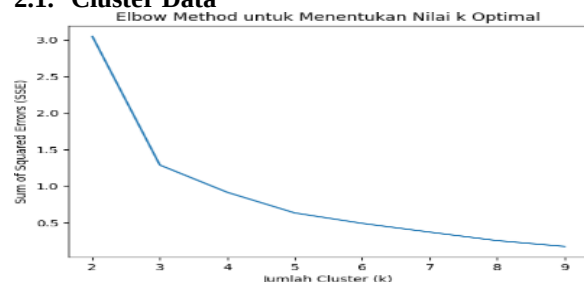
Pendekatan konvensional dalam menangani kasus pelecehan seksual seringkali kurang efektif karena tidak didasarkan pada analisis data yang mendalam[1]

Oleh karena itu, diperlukan metode yang lebih ilmiah dan berbasis data untuk memahami pola dan karakteristik dari kasus-kasus pelecehan seksual anak di bawah umur. Salah satu pendekatan yang dapat digunakan adalah teknik clustering, khususnya algoritma k-means, yang mampu mengelompokkan data berdasarkan kesamaan karakteristik sehingga pola tersembunyi dapat terungkap[2].

Karena kesederhanaan dan efisiensi, algoritma k-means banyak digunakan dalam pengelompokan[3].

Ini karena lebih terukur dan efisien untuk pengolahan banyak objek 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[4]

2.1. Cluster Data



Gambar 1. Kurva Elbow Method

Dari grafik pada gambar 1, dapat dilihat bahwa peneliti memilih jumlah cluster $k=3$ berdasarkan kurva Elbow Method. Hasil clustering ini direpresentasikan dalam bentuk label cluster yang ditambahkan ke dalam dataset asli. Setiap cluster mencerminkan kelompok kabupaten yang memiliki pola serupa dalam distribusi kasus pelecehan seksual[5].

2.2. Data Anak-anak

Tabel 1. Dataset

Kabupaten	2021	2022	2023	2024
Buton	10	7	21	3
Muna	2	16	10	2
Kolaka	22	15	21	5
Konawe Selatan	9	46	43	10
Baubau	25	25	35	19
Kendari	33	42	47	13
Bombana	7	14	14	4
Wakatobi	7	6	7	1
Konawe Utara	7	5	7	1
Konawe	8	15	18	7
Kolaka Utara	8	17	20	13
Kolaka Timur	2	0	3	6
Muna Barat	5	17	25	0
Buton Tengah	3	11	14	3
Buton Selatan	9	15	19	10
Buton Utara	11	6	24	5

Dari Tabel 1 dapat dilihat bahwa data yang digunakan diperoleh dari laporan resmi yang mencatat kasus pelecehan seksual anak di bawah umur di Sulawesi Tenggara selama periode 2021-2024. Data ini mencakup informasi jumlah kasus per kabupaten setiap tahunnya, disusun dalam format CSV dengan kolom-kolom seperti 'kabupaten', 'thn2021', 'thn2022', 'thn2023', dan 'thn2024'[6].

2.3. Rumus K-Means

Menentukan jarak yang paling pendek dari setiap data dengan menggunakan rumus jarak geometri (Euclidean Distance)[7]:

$$d(x_i, \mu_j) = \sqrt{\sum (x_i - \mu_j)^2} \quad (1)$$

Keterangan:

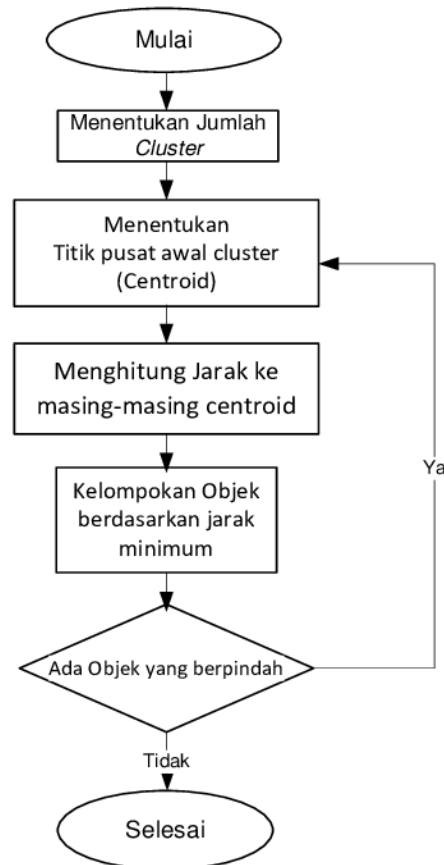
x_i : data kriteria

μ_j : centroid pada cluster ke-j

2.4. Flowchart

Pada gambar 2. Ini bahwa Flowchart menunjukkan tahapan algoritma K-means untuk menganalisis data kasus pelecehan seksual anak. Prosesnya dimulai dengan menentukan jumlah kelompok (cluster)[8].

Kemudian mencari titik pusat awal. Selanjutnya, sistem akan menghitung jarak setiap data ke titik pusat dan mengelompokkannya berdasarkan jarak terdekat. Proses ini akan terus berulang sampai tidak ada lagi data yang berpindah kelompok[9].



Gambar 2. Flowchart K-Means

3. METODE PENELITIAN

3.1. Metode K-Means Clustering

K-Means adalah algoritma clustering berbasis centroid yang digunakan untuk mengelompokkan data ke dalam kkk cluster berdasarkan kemiripan atribut[10].

Algoritma ini bekerja dengan menentukan pusat cluster secara acak, menghitung jarak terdekat menggunakan rumus Euclidean, mengklasifikasikan data, memperbarui pusat cluster, dan mengulangi proses hingga tidak ada perubahan anggota cluster. Algoritma ini banyak digunakan karena sederhana dan efektif. Adapun tahapan dalam penelitian ini:

- Dataset mencatat kasus pelecehan seksual anak di bawah umur di Sulawesi Tenggara periode 2021–2024, mencakup jumlah kasus per kabupaten per tahun. Data disusun dalam format CSV.
- Data diimpor menggunakan Python dan diproses dengan library pandas. Proses meliputi normalisasi menggunakan StandardScaler untuk menyelaraskan skala variabel sebelum dilakukan clustering.
- Penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow untuk mencari titik optimal berdasarkan inersia (SSE).
- Algoritma K-Means diterapkan pada data yang telah dinormalisasi. Cluster hasil analisis ditambahkan ke dataset asli untuk memetakan kasus ke dalam kelompok tertentu.

- e. Hasil clustering divisualisasikan menggunakan scatter plot untuk pola distribusi antar tahun dan pie chart untuk distribusi total kasus di setiap cluster, memudahkan interpretasi pola kasus pelecehan di wilayah tersebut.

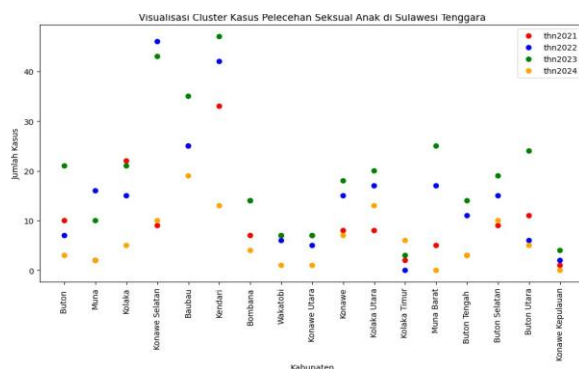
4. HASIL DAN PEMBAHASAN

4.1. Menentukan Jumlah Klaster

Setelah menerapkan algoritma k-means pada data Kasus kekerasan seksual terhadap anak di bawah umur di Sulawesi Tenggara, peneliti mendapatkan hasil clustering dengan menggunakan jumlah cluster optimal yang telah ditentukan. Peneliti memilih jumlah cluster $k=3$ berdasarkan kurva Elbow Method. Hasil clustering ini direpresentasikan dalam bentuk label cluster yang ditambahkan ke dalam dataset asli. Setiap cluster mencerminkan kelompok kabupaten yang memiliki pola serupa dalam distribusi kasus pelecehan seksual. Bisa dilihat pada gambar 1

4.2. Scatter Plot

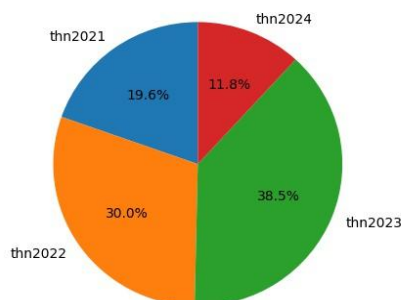
Scatter Plot digunakan untuk menampilkan distribusi jumlah kasus per tahun di setiap kabupaten. Setiap tahun direpresentasikan dengan warna yang berbeda, sehingga memudahkan untuk melihat perubahan seiring waktu dan pola spasialnya. Berikut adalah gambar visualisasi menggunakan Scatter Plot :



Gambar 3. Distribusi jumlah kasus per tahun

4.3. Pie Chart Total Kasus

Total Kasus Pelecehan Seksual Anak di Sulawesi Tenggara (2021-2024)



Gambar 4. Total Kasus

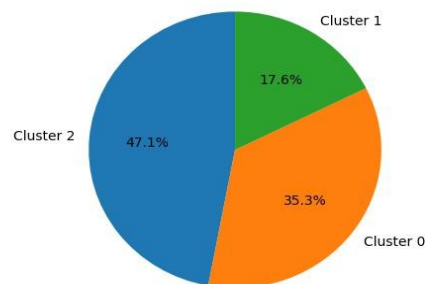
Pie Chart (diagram pie) disebut juga diagram atau grafik lingkaran. Diagram Pie digunakan untuk

menunjukkan ukuran sebagai pembanding data. Pie Chart Total Kasus menunjukkan persentase kasus pelecehan seksual anak per tahun dari total kasus dalam periode 2021-2024. Ini memberikan gambaran komparatif tentang frekuensi kasus pada setiap tahunnya.

4.4. Pie Chart Distribusi Cluster

Pie Chart Distribusi Cluster menampilkan persentase jumlah kabupaten yang masuk ke dalam masing-masing cluster. Diagram ini membantu dalam memahami proporsi relatif dari setiap cluster terhadap total kasus.

Distribusi Cluster Kasus Pelecehan Seksual Anak di Sulawesi Tenggara



Gambar 5. Persentase jumlah kabupaten dalam cluster

4.5. Hasil Clustering

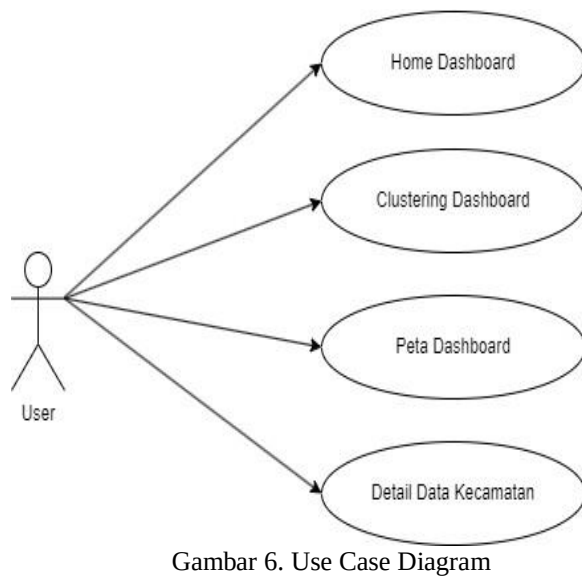
Hasil ini menunjukkan bahwa pembagian wilayah menjadi 3 cluster berbeda berdasarkan karakteristik kasus. Terlihat adanya pola geografis dimana wilayah-wilayah yang berdekatan cenderung masuk dalam cluster yang sama.

Tabel 2. Hasil Clustering

Cluster 0	Cluster 2	Cluster 1
Muna	Buton	Konawe Selatan
Wakatobi	Kolaka	Baubau
Konawe Utara	Bombana	Kendari
Kolaka Timur	Konawe	
Buton Tengah	Kolaka Utara	
Konawe Kepulauan	Muna Barat	
	Buton Selatan	
	Buton Utara	

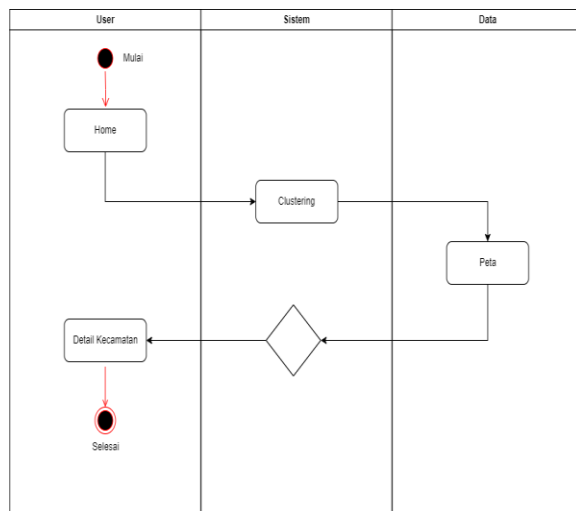
- Cluster 0 memiliki karakteristik jumlah kasus pelecehan seksual yang cenderung lebih rendah dibandingkan dengan klaster lainnya.
- Cluster 1 menunjukkan bahwa daerah-daerah dengan tren peningkatan signifikan dalam kasus pelecehan seksual, atau bisa jadi terdapat faktor-faktor spesifik dalam konteks sosial, budaya, atau kebijakan daerah yang berpengaruh pada kasus-kasus ini.
- Cluster 2 dapat diinterpretasikan sebagai kelompok kabupaten dengan jumlah kasus pelecehan seksual yang cenderung lebih tinggi

4.6. Unified Modeling Language



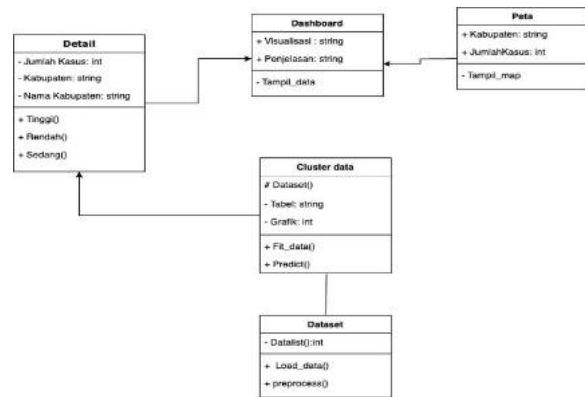
Gambar 6. Use Case Diagram

Dari gambar 6. ini bahwa use case diagram mengilustrasikan fungsionalitas yang diinginkan dari sebuah sistem. Use case bertujuan untuk memodelkan interaksi antara aktor dan sistem. Aktor adalah entitas manusia yang terlibat dalam sistem untuk melakukan tugas-tugas tertentu.



Gambar 7. Activity Diagram

Dari gambar 7. Ini bahwa Activity Diagram menggambarkan aliran kerja atau aktivitas dalam sistem yang akan dieksekusi. Diagram ini juga digunakan untuk mendefinisikan atau mengelompokkan proses dalam sistem. Gambar aktivitas memiliki komponen Activity Diagram menggambarkan aliran kerja atau aktivitas dalam sistem yang akan dieksekusi.



Gambar 7. Class Diagram

Dari gambar 7. bahwa class diagram adalah gabungan dari beberapa kelas dan hubungannya. Setiap kelas direpresentasikan dalam bentuk persegi dengan nama kelas di bagian atas, diikuti oleh atribut kelas di bagian tengah, dan metode kelas di bagian bawah

4.7. Tampilan Interface

Dalam rangka mendukung penelitian ini, saya juga mengembangkan sebuah mockup berbasis web yang dirancang untuk memvisualisasikan hasil clustering kasus pelecehan seksual anak di bawah umur di Sulawesi Tenggara

4.8. Home

Pada halaman home, pengguna disajikan dengan dua bagian utama visualisasi data asli per tahun dan penjelasan tentang proses clustering yang digunakan.



Gambar 8. Tampilan home website

4.9. Clustering

Halaman clustering pada website dirancang untuk menyajikan hasil analisis clustering yang dilakukan menggunakan algoritma k-means. Halaman ini bertujuan untuk memberikan informasi terperinci mengenai pengelompokan distrik di Sulawesi Tenggara berdasarkan jumlah kasus pelecehan seksual anak di bawah umur.



Gambar 9. Tampilan hasil clustering

4.10. Peta

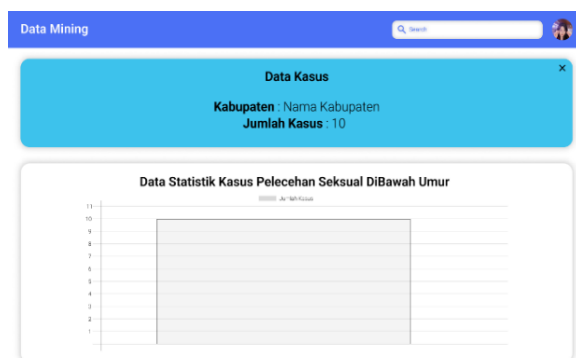
Halaman peta pada website ini dirancang untuk memberikan visualisasi geografis dari jumlah kasus pelecehan seksual anak di bawah umur di Sulawesi Tenggara. Dengan menggunakan peta interaktif.



Gambar 10. Tampilan peta sebaran kasus

4.11. Detail Kecamatan

Halaman detail kecamatan pada website ini dirancang untuk memberikan informasi mendalam mengenai jumlah kasus pelecehan seksual anak di bawah umur di tingkat kecamatan di Sulawesi Tenggara. Halaman ini berfungsi sebagai sumber informasi yang rinci untuk setiap kecamatan, memungkinkan pengguna untuk mendapatkan ilustrasi yang lebih jelas tentang deskripsi di masing-masing daerah.



Gambar 11. Tampilan data kasus tingkat kecamatan

5. KESIMPULAN DAN SARAN

Penelitian ini menggunakan algoritma K-Means untuk mengelompokkan data kasus pelecehan seksual

anak di Sulawesi Tenggara selama periode 2021-2024. Berdasarkan hasil clustering, ditemukan nilai $k=3$ yang mencerminkan pola distribusi kasus di berbagai kabupaten. Klaster pertama (klaster 0) terdiri dari kabupaten Muna, Wakatobi, Konawe Utara, Kolaka Timur, Buton Tengah, dan Konawe Kepulauan dengan jumlah kasus yang cenderung rendah. Klaster kedua (klaster 1) mencakup kabupaten Konawe Selatan, Baubau, dan Kendari dengan peningkatan signifikan dalam kasus, dan klaster ketiga (klaster 2) meliputi kabupaten Buton, Kolaka, Bombana, Konawe, Kolaka Utara, Muna Barat, Buton Selatan, dan Buton Utara dengan jumlah kasus yang sedang hingga tinggi. Analisis lebih lanjut mengungkap bahwa faktor sosial-ekonomi dan efektivitas program perlindungan anak berperan dalam perbedaan prevalensi kasus di masing-masing klaster. Hasil visualisasi dengan scatter plot dan pie chart membantu dalam memahami distribusi kasus dan proporsi relatif dari setiap klaster. Metode Elbow digunakan untuk menentukan jumlah cluster optimal, yang kemudian diaplikasikan dalam algoritma K-Means.

Saran untuk penelitian selanjutnya adalah menyelidiki lebih lanjut pengaruh faktor sosial ekonomi terhadap prevalensi insiden pelecehan seksual terhadap anak dan menggunakan algoritma lain seperti regresi logistik dan pohon keputusan untuk memprediksi potensi insiden. Selain itu, penelitian di masa depan dapat membandingkan efektivitas program perlindungan anak di wilayah yang kebijakannya kuat dan lemah. Wawasan yang lebih mendalam juga dapat diperoleh dengan menggunakan analisis spasial dengan GIS untuk memetakan sebaran kasus di tingkat mikro, seperti desa atau kelurahan. Menilai dampak dari langkah-langkah perlindungan anak yang dilaksanakan di setiap kabupaten dan mengadopsi pendekatan multivariat untuk memahami faktor-faktor penyebab juga akan membantu mengembangkan solusi yang lebih efektif.

DAFTAR PUSTAKA

- [1] R. Fauziah and A. I. Purnamasari, "Implementasi Algoritma K-Means pada Kasus Kekerasan Anak dan Perempuan Berdasarkan Usia," *Hello World Jurnal Ilmu Komputer*, vol. 2, no. 1, pp. 34–41, Mar. 2023, doi: 10.56211/helloworld.v2i1.232.
- [2] R. R. Putra and C. Wadisman, "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K Means," *INTECOMS: Journal of Information Technology and Computer Science*, vol. 1, no. 1, pp. 72–77, Mar. 2018, doi: 10.31539/intecom.v1i1.141.
- [3] H. Jurnal, "JURNAL MANAJEMEN DAN TEKNIK INFORMATIKA," *JUMANTAKA*, vol. 02, p. 1, 2018.

-
- [4] W. Sudrajat, I. Cholid, and J. Petrus, "Wahyu Sudrajat et al, Penerapan Algoritma K-Means Untuk"
- [5] F. Sutomo et al., "OPTIMIZATION OF THE K-NEAREST NEIGHBORS ALGORITHM USING THE ELBOW METHOD ON STROKE PREDICTION," *Jurnal Teknik Informatika (JUTIF)*, vol. 4, no. 1, 2023, doi: 10.20884/1.jutif.2023.4.1.839.
- [6] P. Agustia Rahayuningsih, R. Maulana, P. Studi Komputerisasi Akuntansi, A. BSI Pontianak, and J. Abdurahman Saleh, "Analisis Perbandingan Algoritma Klasifikasi Data Mining Untuk Dataset Blogger Dengan Rapid Miner," vol. VI, no. 1, 2018.
- [7] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," vol. 15, no. 2.
- [8] D. Q. Zeebaree, H. Haron, A. Mohsin Abdulazeez, and S. R. M. Zeebaree, "Combination of K-means clustering with Genetic Algorithm: A review," 2017. [Online]. Available: <http://www.ripublication.com>
- [9] R. Rosaly, A. Prasetyo, and M. Kom, "Pengertian Flowchart Beserta Fungsi dan Simbol-simbol Flowchart yang Paling Umum Digunakan."
- [10] S. Handoko, F. Fauziah, and E. T. E. Handayani, "IMPLEMENTASI DATA MINING UNTUK MENENTUKAN TINGKAT PENJUALAN PAKET DATA TELKOMSEL MENGGUNAKAN METODE K-MEANS CLUSTERING," *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. 25, no. 1, pp. 76–88, 2020, doi: 10.35760/tr.2020.v25i1.2677