JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9520
Hal 451-461
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Perbandingan Metode Euclidean dan Manhattan Distance dalam Implementasi Algoritma K-Means dan K-Medoid pada Pengelompokkan Faktor Dominan Perceraian di Kabupaten Bojonegoro

Elok Salma Nabila*, Ifnu Wisma Dwi Prastya, Ita Aristia Sa'ida

Sains dan Teknologi, Teknik Informatika, Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro, Indonesia
Email: [1*]eloksalma2019@gmail.com, [2]ifnuprastya@unugiri.ac.id, [3,] itaaristia@unugiri.ac.id
Email Penulis Korespondensi: eloksalma2019@gmail.com
Submitted **13-01-2026;** Accepted **24-02-2026;** Published **28-02-2026**

Abstrak

Tingkat Perceraian di Kabupaten Bojonegoro terus menunjukkan peningkatan dan dipicu oleh berbagai faktor sosial seperti perselisihan, tekanan ekonomi, dan ketidakharmonisan dalam rumah tangga, sehingga diperlukan analisis yang mampu memetakan faktor dominan dan non-dominan secara lebih terarah. Sehingga, penelitian ini bertujuan untuk mengelompokkan faktor penyebab perceraian di Kabupaten Bojonegoro pada periode 2021–2023 serta menentukan metode *clustering* yang paling optimal. Pada penelitian ini menggunakan algoritma K-Means dan K-Medoids dengan jarak Euclidean dan Manhattan pada data asli dan data yang dinormalisasi menggunakan Min–Max Scaler, dengan evaluasi Silhouette Score. Hasil menunjukkan bahwa normalisasi data meningkatkan kualitas klaster, dan K-Means dengan jarak Manhattan pada data ternormalisasi memberikan nilai terbaik, yaitu Silhouette Score 0,849547. Analisis perpindahan klaster menunjukkan bahwa pola pengelompokan relatif konsisten antar tahun, dengan faktor perselisihan terus-menerus muncul sebagai faktor dominan sementara faktor-faktor lain berada pada klaster non-dominan dengan pola yang serupa. Penelitian ini menunjukkan bahwa K-Means dengan jarak Manhattan pada data ternormalisasi lebih efektif untuk mengelompokkan faktor perceraian.Temuan ini dapat menjadi dasar metodologis bagi pemerintah daerah dalam merumuskan kebijakan dan penanganan masalah sosial berbasis data.

Kata Kunci: Perceraian; Clustering; K-Means; K-Medoids; Silhouette Score; Min–Max Scaler

Abstract

The divorce rate in Bojonegoro Regency continues to increase, driven by various social factors such as constant disputes, economic pressure, and household disharmony. Consequently, an analysis is required to map dominant and non-dominant factors more effectively. This study aims to group the factors causing divorce in Bojonegoro Regency for the 2021–2023 period and determine the most optimal clustering method. The research utilizes K-Means and K-Medoids algorithms with Euclidean and Manhattan distance metrics applied to both raw data and data normalized using the Min–Max Scaler, evaluated via the Silhouette Score. The results indicate that data normalization improves cluster quality, and K-Means with Manhattan distance on normalized data achieves the best performance, yielding a Silhouette Score of 0.849547. Cluster displacement analysis reveals that the grouping patterns remain relatively consistent across years, with "constant disputes" consistently emerging as the dominant factor, while other factors remain in the non-dominant cluster with similar patterns. This study demonstrates that K-Means with Manhattan distance on normalized data is more effective for clustering divorce factors. These findings can serve as a methodological foundation for the local government in formulating data-driven social policies and interventions.

Keywords: Divorce; Clustering; K-Means; K-Medoids; Silhouette Score; Min–Max Scaler

1. PENDAHULUAN

Perceraian merupakan fenomena sosial yang mempengaruhi stabilitas keluarga dan struktur sosial masyarakat. Jumlah angka perceraian yang terus meningkat tidak hanya berdampak pada kesehatan psikologis anggota keluarga tetapi juga dapat berdampak pada perubahan sosial yang meningkatkan tanggung jawab ekonomi pada *single parent* dan kemungkinan munculnya konflik sosial di lingkungan mereka [1]. Faktor yang mempengaruhi perceraian diantaranya seperti kekerasan dalam rumah tangga, perselisihan terus menerus,dan ekonomi [2]. Dengan demikian, identifikasi secara menyeluruh terhadap faktor-faktor yang melatarbelakangi terjadinya perceraian sangat diperlukan agar dapat dilakukan penanganan dan pencegahan secara tepat [3].

Faktor yang mempengaruhi perceraian sangatlah beragam diantaranya faktor *demografis*, seperti usia pernikahan yang terlalu muda dan tingkat Pendidikan yang rendah. Faktor sosial juga menyumbang tingkat perceraian seperti buruknya komunikasi antar pasangan, kekerasan dalam rumah tangga, serta faktor ekonomi yang meliputi tekanan finansial dan pendapatan. Dengan beragamnya faktor yang menjadi penyebab perceraian maka sangat diperlukan analisis yang sistematis untuk menentukan faktor dominan dan non dominan yang menyebabkan terjadinya suatu perceraian, supaya langkah-langkah dalam menentukan pencegahan dan penanganan dapat diambil secara tepat.

Data mining merupakan pendekatan analitis yang digunakan untuk memahami struktur dan pola tersembunyi dalam kumpulan data berukuran besar. Data mining digunakan untuk menemukan pola dalam kumpulan data mentah berukuran besar dengan memanfaatkan teknik *machine learning* untuk mengekstraksi pengetahuan yang sebelumnya tidak terlihat [4]. Metode klasterisasi, khususnya algoritma K-Means dan K-Medoids, banyak digunakan karena efektif dalam membentuk kelompok data yang memiliki tingkat kemiripan tinggi meskipun strukturnya kompleks . Meskipun ada perbedaan yang signifikan diantara kedua algoritma ini, tetapi sama-sama termasuk metode *partitional clustering* karena membagi data ke dalam beberapa partisi berdasarkan kemiripan karakteristik meskipun dengan mekanisme berbeda dalam menentukan pusat klaster [5]. K-Means menghitung pusat klaster menggunakan nilai rata-rata, sehingga bekerja cepat namun sensitif terhadap

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9520
Hal 451-461
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

data pencilan yang dapat menggeser posisi centroid dan menurunkan kualitas klaster [6]. Disisi lain, algoritma K-Medoids menentukan pusat klaster dengan menggunakan objek asli yang mempresentasikan kelompoknya sehingga nilai exstrem tidak banyak mempengaruhi pusat klaster. Oleh karena itu, K-Medoids sering dianggap lebih akurat saat digunakan pada data sosial yang tidak seragam seperti data faktor perceraian dengan variasi nilai yang cukup tinggi [7]. Beberapa studi juga menegaskan bahwa K-Medoids menunjukkan ketahanan yang lebih baik terhadap outlier dan noise, sehingga mampu menghasilkan hasil clustering yang lebih stabil dibandingkan metode berbasis rata-rata [8].

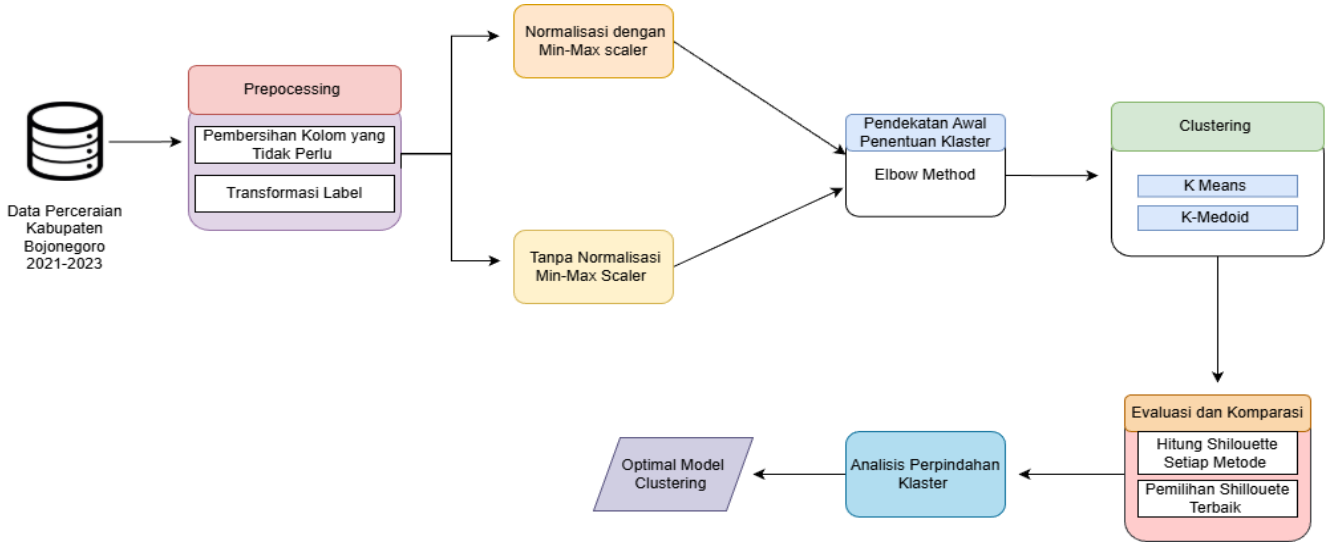
Beberapa Penelitian sebelumnya telah menggunakan metode K-Means untuk menganalisis kasus perceraian di Indonesia, tetapi masih sangat terbatas penelitian yang membandingkan K-Means dengan K-Medoid di konteks sosial yang kompleks, seperti penyebab perceraian. Data sosial sering kali mengandung ketidakpastian dan nilai pencilan, yang dapat membuat analisis K-Means kurang efektif. Hal ini didukung oleh penelitian Yanti (2021) dengan menggunakan K-Means untuk kasus perceraian di jambi dalam kesimpulannya ia menyarankan agar di penelitian selanjutnya dapat membandingkan atau menggabungkan dua metode atau lebih untuk mendapatkan hasil yang lebih akurat [9]. Di sisi lain, penelitian Tampubolon et al. (2021) memang membandingkan K-Means dan K-Medoids, tetapi dilakukan pada data sosial yang berbeda sehingga belum memberikan gambaran spesifik terkait faktor penyebab perceraian [10]. Selain itu, studi Zahro (2024) di Bojonegoro hanya menggunakan K-Means tanpa metode pembanding maupun evaluasi kualitas klaster yang terukur [11]. Dari temuan tersebut, terlihat bahwa belum ada penelitian yang secara langsung membandingkan kedua algoritma ini pada data faktor penyebab perceraian, dua skenario pengolahan data serta penggunaan dua metrik jarak yaitu Euclidean dan Manhattan dengan penilaian kualitas klaster menggunakan Silhouette Score. Oleh karena itu, penelitian ini menyajikan analisis perbandingan K-Means dan K-Medoids dengan menggunakan 2 metrik jarak yaitu Euclidean dan Manhattan pada data perceraian di Kabupaten Bojonegoro dengan pendekatan evaluasi Shilhouette Score dan juga menggunakan 2 skenario data.

Penelitian ini bertujuan untuk mengelompokkan faktor-faktor penyebab kasus perceraian di Kabupaten Bojonegoro dan mengidentifikasi faktor yang mendominasi penyebab kasus perceraian pada rentang tahun 2021 hingga 2023 menggunakan pendekatan *clustering* yaitu dengan menggunakan algoritma K-Means dan K-Medoids. Menganalisis perbedaan klaster yang dihasilkan oleh kedua metode tersebut berdasarkan kemiripan karakteristik data, serta membandingkan kualitas klaster yang dihasilkan untuk menentukan algoritma paling efektif yang mereprentasikan pola perceraian di Kabupaten Bojonegoro menggunakan metrik evaluasi yang tepat seperti Shillouete Score. Selain itu, penelitian ini menggunakan 2 skenario pengolahan data, yaitu data yang telah dinormalisasi dengan Min-Max Scaler, dan data tanpa normalisasi untuk menganalisis pengaruh hasil klasterisasi yang dihasilkan.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian disusun secara berurutan dan saling terhubung untuk memastikan proses penelitian berjalan terstruktur sesuai tujuan, mulai dari pengolahan data hingga evaluasi model. Alur tersebut disajikan pada gambar berikut.



Gambar 1. Tahapan Penelitian

Pada gambar 1 menyajikan beberapa rangkaian langkah dari penelitian yang tersusun dimulai dari pengumpulan data perceraian hingga proses evaluasi. Dari setiap bagian alur tersebut mendukung terbentuknya model *clustering* yang tepat dalam memetakan pola perceraian yang terjadi di Kabupaten Bojonegoro. Berikut penjelasan lengkap dari masing masing tahapan.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9520
Hal 451-461
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

2.2 Pengumpulan Data

Pengumpulan data adalah dasar utama proses analisis karena kualitas hasil akhir sangat bergantung pada tahap pembersihan dan perancangan data yang dilakukan sebelumnya [12]. Sumber data yang dijadikan penelitian berasal dari studi terdahulu yang disusun oleh Zahro (2024) serta dokumen resmi Pengadilan Agama Kabupaten Bojonegoro pada tahun 2021-2023. Pengumpulan ini dilakukan untuk membangun dataset yang valid mengenai faktor-faktor yang mempengaruhi terjadinya perceraian. Data tersebut berisi beberapa alasan perceraian dari pihak penggugat dengan variabel seperti: meninggalkan salah satu pihak, ekonomi, kekerasan dalam rumah tangga (KDRT), perselisihan terus menerus, faktor ekonomi, poligami, dan beberapa faktor lainnya. Seluruh data disajikan dengan tabel sehingga bisa dilakukan pengolahan data dengan teknik data mining. Selain itu, pengumpulan data ini juga mencakup penyelarasan data selama bertahun tahun. Sehingga dataset ini memberi gambaran tentang faktor perceraian selama tiga tahun.

2.3 *Preprocessing* Data

Preprocessing diperlukan agar data layak dianalisis, karena kualitas data mentah sangat memengaruhi hasil data mining [13]. Pada tahap ini dilakukan beberapa langkah penting, yaitu:

a. Menghapus kolom yang tidak perlu

Proses *prepocessing* dimulai dengan menghapus kolom jumlah yang terdapat pada dataset. Kolom ini menunjukkan total keseluruhan kasus perceraian. Karena variabel faktor penyebab perceraian yang menjadi fokus utama dalam penelitian ini, maka kolom jumlah tidak diperlukan dan dihapus untuk menjaga efisiensi dan konsistensi pada data. Jika kolom jumlah tetap disertakan, nilai tersebut akan membuat hasil klaster bias.

b. *Transformasi* label

Dataset perceraian memiliki struktur awal label dibagian baris pada tabel yang berisi kategori faktor penyebab perceraian. sementara itu, label informasi bulan perceraian tercantum pada bagian kolom tabel. Struktur ini kurang sesuai untuk proses *clustering* pada penelitian ini, karena studi ini bertujuan untuk mengidentifikasi pola faktor dominan dan non-dominan, bukan mengelompokkan berdasarkan bulan. Oleh karena itu, tahap *transformasi* label dilakukan untuk menata ulang struktur dari dataset perceraian, dimana faktor perceraian dijadikan sebagai atribut utama. Sementara itu, informasi bulan pada kasus perceraian tetap menjadi informasi pendukung, tetapi tidak menjadi komponen penting.

2.4 Normalisasi Data

Normalisasi Data dilakukan untuk menyeragamkan skala setiap atribut agar perbedaan nilai antar fitur tidak menimbulkan bias [14]. Pada penelitian ini digunakan dua skenario normalisasi, yaitu:

a. Data dengan Normalisasi

Pada skenario ini, data dinormalisasi menggunakan Min–Max Scaler agar setiap atribut berada pada rentang yang sama. Nilai normalisasi dihitung dengan rumus:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Rumus ini digunakan untuk menghitung hasil normalisasi x' dengan cara mengurangkan nilai dari data asli x dengan nilai minimum atribut (*Xmin*), setelah itu membaginya dengan selisih antara nilai maximum (*Xmax)* dan nilai minimum atribut.

b. Data Tanpa Normalisasi

Pada skenario kedua, data digunakan dalam bentuk aslinya tanpa penyetaraan skala. Skenario ini bertujuan melihat bagaimana algoritma bekerja ketika setiap atribut memiliki rentang nilai berbeda. Hasil klasterisasi dari kedua skenario kemudian dibandingkan untuk menilai pengaruh normalisasi terhadap kualitas klaster.

2.5 Pendekatan Awal Penentuan Klaster dengan Elbow Method

Elbow Method digunakan untuk memperkirakan jumlah klaster yang paling sesuai dengan melihat perubahan *within-cluster sum of squares* ketika jumlah klaster ditambah. Nilai keragaman biasanya turun tajam di awal lalu melambat setelah mencapai titik yang membentuk "siku". Titik tersebut menandai jumlah klaster yang tidak lagi memberikan peningkatan signifikan, sehingga Elbow Method sering dijadikan dasar awal dalam menentukan jumlah klaster pada algoritma K-Means [15].

2.6 *Clustering*

Proses utama dalam penelitian ini adalah pada tahap clustering, dimana data perceraian di kelompokkan berdasarkan kemiripan karakteristik [16]. Tahap *clustering* dimulai dengan menggunakan dua algoritma yaitu K-Means dan K-Medoids. Metode K-Means menggunakan pendekatan rata-rata (*centroid*) sebagai pusat klaster dengan tujuan meminimalkan jumlah jarak kuadrat dari setiap titik ke pusat klaster [17][18]. Sementara itu, K-Medoids menggunakan titik data aktual (*medoid*) sebagai pusat klaster sehingga lebih tahan terhadap pencilan (*outlier*) dan nilai ekstrem [19].

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9520
Hal 451-461
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

a. Proses *Clustering* Menggunakan K-Means

Pada proses *clustering* yang pertama yaitu menggunakan Metode K-Means. Dalam penelitian ini, dua metrik digunakan untuk mengukur jarak yaitu jarak Euclidean Distance dan Manhattan Distance.

1. Euclidean Distance

$$d(X_i, C_k) = \sqrt{\sum_{j=1}^{n} \left(x_{ij} - c_{kj}\right)^2} \quad (2)$$

Rumus digunakan untuk menghitung jarak antara ***Xi*** dan pusat klaster ***Ck*** dengan menghitung nilai setiap atribut. Hasilnya menunjukkan seberapa dekat dengan klaster dan jarak terkecil menunjukkan bahwa data tersebut paling sesuai untuk masuk ke klaster tersebut.

2. Manhattan Distance

$$d(X_i, C_k) = \sum_{j=1}^{n} \left|x_{ij} - c_{kj}\right| \quad (3)$$

Nilai jarak ditentukan oleh total selisih absolut seluruh atribut antara data ***Xi*** dan centroid ***Ck***.

Algoritma K-Means bekerja secara iteratif yang dimulai dengan menentukan sejumlah k pusat klaster (centroid) secara acak. Selanjutnya, setiap data dihitung jaraknya terhadap centroid tersebut untuk dimasukkan ke dalam kelompok dengan jarak terdekat. Setelah semua data terbagi, posisi centroid diperbarui berdasarkan nilai rata-rata dari anggota baru di masing-masing klaster [20]. Siklus ini terus berulang hingga posisi centroid tidak lagi berubah, yang menandakan bahwa hasil pengelompokan telah stabil [21].

b. Proses Clustering Menggunakan K-Medoid

Pada proses clustering ini juga digunakan metode K-Medoids. K-Medoids merupakan metode pengelompokan berbasis partisi yang menggunakan pusat klaster berupa medoid, yaitu titik data aktual yang paling mewakili anggota klaster dan menggunakan objek data asli sebagai pusat klaster [22] [23]. Pemilihan medoid awal dilakukan dengan memilih dua data asli secara acak dari dataset. Setelah itu, kedekatan setiap data terhadap medoid dihitung menggunakan dua jenis jarak.

1. Euclidean Distance

$$d_E(X_i, M_k) = \sqrt{\sum_{j=1}^{n} \left(x_{ij} - m_{kj}\right)^2} \quad (4)$$

Rumus digunakan untuk menghitung jarak Euclidean antara data ke-*i* (***Xi***) dan pusat klaster ke-*k* (***Mk***) dengan menjumlahkan kuadrat selisih setiap atribut, kemudian diambil akar kuadratnya

2. Manhattan Distance

$$d_M(X_i, M_k) = \sum_{j=1}^{n} \left|x_{ij} - m_{kj}\right| \quad (5)$$

Rumus menghitung jarak Manhattan antara data (***Xi)*** dan pusat klaster (***Mk)*** sebagai penjumlahan nilai absolut selisih setiap atribut *xij* terhadap *mkj* untuk seluruh *n* atribut, tanpa melibatkan operasi pemangkatan maupun akar

Berdasarkan hasil penghitungan jarak, total cost dihitung dengan menjumlahkan seluruh jarak data terhadap medoid pada masing-masing klaster, baik menggunakan Euclidean maupun Manhattan.

1. Cost Euclidean Distance

$$\text{Cost}_E = \sum_{k=1}^{K} \sum_{X_i \in C_k} d_E(X_i, M_k) \quad (6)$$

2. Cost Manhattan Distance

$$\text{Cost}_M = \sum_{k=1}^{K} \sum_{X_i \in C_k} d_M(X_i, M_k) \quad (7)$$

Rumus menghitung total cost sebagai ukuran kualitas pengelompokan, yaitu dengan menjumlahkan jarak antara setiap data ***Xi*** dan pusat klaster ***Mk*** pada masing-masing klaster ***Ck*** Perhitungan total cost dilakukan untuk seluruh klaster.

Nilai cost ini digunakan untuk menilai kualitas pemilihan medoid. Apabila ditemukan medoid lain yang menghasilkan total cost lebih kecil, medoid akan diganti melalui proses swap [24], dan seluruh perhitungan diulang kembali. Proses ini berlangsung hingga tidak ada penurunan cost, sehingga medoid yang diperoleh dianggap paling optimal.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9520
Hal 451-461
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

c. Evaluasi dan Komparasi

Pada tahap ini, nilai silhouette score dihitung untuk setiap metode klasterisasi pada masing-masing tahun dan digunakan sebagai dasar pembandingan kinerja. Silhouette score dipilih karena bersifat umum, tidak bergantung pada algoritma tertentu, serta dapat digunakan pada berbagai metrik jarak [25]. Secara konsep silhouette score dirumuskan sebagai berikut:

$$s = \frac{b - a}{\max(a, b)} \tag{8}$$

Dengan a menyatakan rata rata jarak antar objek dalam klaster yang sama, dan b menyatakan rata-rata jarak objek terhadap klaster terdekat lainnya. Metode dengan nilai silhouette score tertinggi ditetapkan sebagai metode paling optimal.

d. Analisis Perpindahan Klaster

Tahapan ini difokuskan pada analisis perpindahan klaster antar tahun dengan menerapkan metode *clustering* yang telah ditetapkan sebagai metode terbaik berdasarkan hasil evaluasi Silhouette Score. Metode tersebut digunakan secara konsisten pada seluruh periode pengamatan, yaitu tahun 2021 hingga 2023, agar hasil pengelompokan yang dihasilkan memiliki dasar yang sama dan dapat dibandingkan secara langsung antarperiode waktu. Analisis dilakukan dengan membentuk klaster untuk setiap tahun pengamatan menggunakan metode terpilih, kemudian menelusuri perubahan keanggotaan klaster yang terjadi dari satu tahun ke tahun berikutnya. Tahapan ini bertujuan untuk mengamati pola pergeseran klaster secara temporal**,** sekaligus menilai stabilitas struktur klaster dan peran variabel dalam memengaruhi perubahan pengelompokan**.**

3. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian dan pembahasan yang diperoleh dari seluruh tahapan penelitian yang telah dilakukan. Pembahasan disusun secara berurutan sesuai dengan alur tahapan penelitian yang telah dijelaskan sebelumnya di bab 2.

3.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data dari Pengadilan Agama Kabupaten Bojonegoro. Data ini memuat informasi kasus perceraian yang terjadi pada periode 2021 hingga 2023. Rincian data disajikan dalam tabel berikut.

Tabel 1. Pengumpulan Data Kasus Perceraian

Bulan	Zina	Mabuk	Madat	Judi	Meninggalkan satu pipihak	Dihukum penjara	Poligami	Kekerasan dalam RT	Cacat badan	Perselihan terus menerus	Kawin paksa	murtad	Ekonomi	Lain lain	Jumlah
Januari	0	0	1	0	21	0	0	6	0	106	5	1	88	0	228
Februari	0	0	0	2	8	0	0	3	0	51	1	0	104	0	169
……	….	….	….	…	…	….	…	….	…	…	…	..	…	…	…
Desember	0	1	0	1	12	0	0	4	0	62	4	0	138	0	222

Berdasarkan tabel 1 data mencakup 14 faktor penyebab faktor perceraian dan informasi bulan terjadinya kasus perceraian pada setiap tahun. Rincian yang disajikan dalam tabel, dimana baris menunjukkan faktor penyebab perceraian serta jumlah kasus dan kolom menunjukkan informasi setiap bulan terjadinya setiap kasus perceraian.

3.2 Hasil *Prepocessing* Data

a. Menghapus kolom tidak perlu

Untuk menghasilkan klaster yang lebih mencerminkan kondisi sesungguhnya, penyusunan dataset dilakukan dengan mengecualikan kolom “Jumlah”. Penghapusan kolom “Jumlah” dilakukan untuk mencegah dominasi nilai total pada perhitungan jarak, sehingga proses pengelompokkan lebih mencerminkan variasi pada atribut faktor penyebab perceraian. Berikut hasil menghapus kolom jumlah disajikan pada tabel berikut,

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9520
Hal 451-461
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Tabel 2. Menghapus Kolom Jumlah

Bulan	Zina	Mabuk	Madat	Judi	Meninggalkan satu pihak	Dihukum penjara	Poligami	Kekerasan dalam RT	Cacat badan	Perselihan terus menerus	Kawin paksa	murtad	Ekonomi	Lain lain
Januari	0	0	1	0	21	0	0	6	0	106	5	1	88	0
Februari	0	0	0	2	8	0	0	3	0	51	1	0	104	0
……	….	….	….	…	…	….	…	….	…	…	…	..	…	….
Desember	0	1	0	1	12	0	0	4	0	62	4	0	138	0

Terlihat pada tabel 2 bahwa dataset setelah proses preprocessing data sudah tidak memuat kolom "Jumlah", sehingga atribut yang tersisa adalah faktor penyebab perceraian perbulan.

b. *Transformasi* Label

Pada tahap *tranformasi* label dilakukan penataan ulang struktur dataset agar sesuai dengan tujuan klasterisasi. Struktur awal data dinilai kurang sesuai karena label pada faktor penyebab perceraian dan informasi bulan pada posisi yang belum mendukung untuk dilakukan analisis faktor dominan dan non dominan. Oleh karena itu, dataset disusun ulang untuk menjadikan faktor penyebab perceraian sebagai atribut utama dan informasi bulan tetap menjadi atribut pendukung agar data mudah untuk diiterpretasikan oleh algoritma klasterisasi. Hasil dari Transformasi Label disajikan pada tabel berikut.

Tabel 3. *Transformasi* Label

faktor	Januari	Februari	Maret	April	Mei	Juni	Juli	Agustus	September	Oktober	November	Desember
Zina	0	0	1	0	21	0	0	6	0	106	5	**1**
Mabuk	0	0	0	2	8	0	0	3	0	51	1	**0**
…..	….	….	….	…	…	….	…	….	…	…	…	**..**
Lain lain	0	1	0	1	12	0	0	4	0	62	4	**0**

Berdasarkan tabel 2 dataset sudah disusun ulang dan menempatkan faktor penyebab perceraian menjadi atribut utama.

3.3 Hasil Tahap Normalisasi Data

a. Data dengan Normalisasi Min-Max Scaler

Pada skenario ini, dataset dinormalisasi menggunakan min-max scaler untuk menyetarakan skala nilai pada setiap atribut sehingga perbedaan rentang antar fitur dapat diminimalkan. Normalisasi ini bertujuan agar setiap atribut memiliki kontribusi yang lebih seimbang. Hasil normalisasi data disajikan pada tabel berikut.

Tabel 4. Hasil Normalisasi Data

faktor	Januari	Februari	Maret	April	Mei	Juni	Juli	Agustus	September	Oktober	November	Desember
Zina	0	0	0	7,353	0	0	0	0	0	0	0	0
Mabuk	0	0	11,976	14,706	24	9,376	5,319	0	0	0	0	7,246
…..	….	….	….	…	…	….	…	….	…	…	…	..
Lain lain	0	0	0	0	0	0	0	0	0	0	0	0

Pada Tabel 4 menunjukkan bahwa setiap atribut telah dinormalisasi menggunakan Min-Max Scaler dan berada pada skala yang sama. Dengan demikian perhitungan jarak pada proses klasterisasi tidak didominasi atribut tertentu.

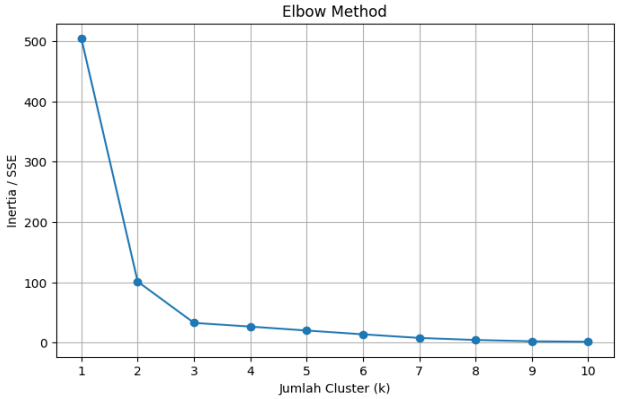
b. Data Tanpa Normalisasi

Data digunakan dalam kondisi sesuai dataset asli setelah melalui tahap *prepocessing*, tanpa dilakukan normalisasi menggunakan Min-Max Scaler. Setiap atribut memiliki rentang nilai yang berbeda sesuai dengan karakteristik awal.

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9520
Hal 451-461
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

3.4 Pendekatan Awal Penentuan Klaster dengan Elbow Method

Dalam penelitian ini, pendekatan awal penentuan klaster dilakukan menggunakan elbow method, dengan mempertimbangkan pola penurunan pada *inertia*. Hasilnya di sajikan pada gambar berikut.



Gambar 2. Grafik Elbow Method

Berdasarkan gambar 2, penurunan *inertia* terlihat paling tajam pada k= 3 dan mulai melandai setelahnya. Dengan demikian, meskipun cluster pada gambar diatas optimal menggunakan k=3, interpretasi dan analisis pada tahap selanjutnya difokuskan pada dua klaster agar fokus analisis lebih terarah dan konsisten dengan tujuan penelitian.

3.5 Hasil Klasterisasi

a. *Clusteri*ng dengan K-Means

Klasterisasi faktor penyebab perceraian dilakukan menggunakan algoritma K-Means Euclidean dan Manhattan pada dua skenario pengolahan data, yaitu tanpa normalisasi dan dengan normalisasi Min-Max Scaler. Proses pengelompokan diterapkan terpisah pada data perceraian tahun 2021-2023 dengan jumlah klaster dua. Pada setiap pengujian, langkah yang dilakukan bersifat sama, yaitu, inisialisasi pusat klaster, perhitungan jarak setiap data ke pusat klaster, penentuan klaster berdasarkan jarak terdekat, pembaruan pusat klaster, serta iterasi hingga konvergen. Tabel 5 menyajikan hasil klasterisasi K-Means menggunakan jarak Euclidean pada data tanpa normalisasi.

Tabel 5. Hasil *Clustering* K-Means Euclidean Tanpa Normalisasi

C1	C2	Jarak Terdekat	Cluster	A(I)	B(I)	S(I)
9.230769231	316.6010897	9.230769231	1	9.230769231	316.6010897	0.9708441647
7.180703308	339.8014126	7.180703308	1	7.180703308	339.8014126	0.9788679416
211.3525392	139.1186544	139.1186544	2	139.1186544	211.3525392	0.341769657
478.0277842	139.1186544	139.1186544	2	139.1186544	139.1186544	0.7089737062
.....	.....	.....	.....	.....	.....	.....
109.3709285	231.8093985	109.3709285	1	109.3709285	231.8093985	0.5281859613
		Rata-Rata Silhouette				**0,8344**

Pada tabel 5 menampilkan hasil *clustering* menggunakan algoritma K-Means dengan menggunakan pendekatan jarak Euclidean pada data tanpa normalisasi. Nilai jarak terhadap klaster pertama (C1) dan klaster kedua (C2) digunakan untuk menentukan keanggotaan klaster berdasarkan jarak yang terdekat. Selanjutnya, pada nilai a(i) dan b(i) dihitung untuk mengukur tingkat kedekatan objek terhadap klaster asal dan klaster alternatif yang kemudian digunakan dalam perhitungan nilai s(i) atau Silhouette Score.

Nilai Silhouette Score yang diperoleh sebagian besar berada pada rentang positif dan mendekati nilai maksimum yaitu dengan rata rata 0,8344, nilai ini mengindikasikan bahwa objek penelitian telah terkelompok dengan baik dan tingkat pemisahan klaster yang jelas. Prosedur perhitungan yang sama juga diterapkan dalam dua metrik jarak, yaitu Euclidean dan Manhattan serta pada dua skenario pengolahan data, yaitu tanpa normalisasi dan normalisasi menggunakan Min-Max Scaler. Berikut tabel 6 menampilkan hasil klaster yang diperoleh dari seluruh skenario pengujian.

Tabel 6. Hasil Cluster K-Means

2021				2022				2023			
Euclidean Tanpa Normalisasi	Euclidean Normalisasi	Manhattan Tanpa Normalisasi	Manhattan Normalisasi	Euclidean Tanpa Normalisasi	Euclidean Normalisasi	Manhattan Tanpa Normalisasi	Manhattan Normalisasi	Euclidean Tanpa Normalisasi	Euclidean Normalisasi	Manhattan Tanpa Normalisasi	Manhattan Normalisasi
1	1	1	1	1	1	1	1	2	1	2	2

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9520
Hal 451-461
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

1	1	1	1	1	1	1	1	2	1	2	2
1	1	1	1	1	1	1	1	2	1	2	2
1	1	1	1	1	1	1	1	2	1	2	2
1	2	1	1	1	1	1	1	2	1	2	2
1	1	1	1	1	1	1	1	2	1	2	2
1	1	1	1	1	1	1	1	2	1	2	2
1	1	1	1	1	1	1	1	2	1	2	2
1	1	1	1	1	1	1	1	2	1	2	2
2	2	2	2	2	2	2	2	1	2	1	1
1	1	1	1	1	1	1	1	2	1	2	2
1	1	1	1	1	1	1	1	2	1	2	2
2	2	2	1	2	1	2	1	1	1	1	2
1	1	1	1	1	1	1	1	2	1	2	2

Pada tabel 6 menampilkan hasil keanggotaan klaster objek penelitian pada periode 2021–2023 yang diperoleh menggunakan algoritma K-Means dengan dua metrik jarak**,** yaitu Euclidean dan Manhattan**,** serta pada dua kondisi data, tanpa normalisasi dan dengan normalisasi Min–Max Scaler**.** Hasil menunjukkan bahwa penggunaan jarak Euclidean tanpa normalisasi pada tahun 2021–2023 menghasilkan struktur klaster yang lebih stabil dibandingkan Euclidean dengan normalisasi, yang menyebabkan terjadinya perpindahan klaster. Sementara itu, jarak Manhattan baik tanpa normalisasi maupun dengan normalisasi cenderung menghasilkan pola klaster yang stabil pada setiap tahun.

b. Hasil *Clustering* dengan K-Medoids

Pengelompokkan data dilakukan menggunakan algoritma K-Medoids dengan pendekatan Euclidean Distance dan Manhattan Distance, dimana setiap objek dialokasikan ke cluster terdekat berdasarkan medoid. Proses ini dilakukan pada 2 skenario data, Berikut tabel 7 hasil *clustering* dengan K-Medoids Euclidean Distance tanpa normalisasi data.

Tabel 7. Hasil *Clustering* K-Medoid Euclidean Tanpa Normalisasi

C1	C2	Jarak Terdekat	Cluster	A(I)	B(I)	S(I)
472.5367711	1.414213562	1.414213562	2	1.414213562	472.5367711	0.0219807066
469.8446552	4.123105626	4.123105626	2	4.123105626	469.8446552	0.0254999487
0	64.33506043	0	1	0	64.33506043	1
4	62.2655603	4	1	4	62.2655603	0.9357590299
.....	.....	.....	.....	.....	.....	.....
0	28.01785145	0	1	0	30.98386677	1
		Rata-Rata Shilouette				0,8176

Tabel 7 menunjukan hasil klasterisasi menggunakan K-Medoids Euclidean tanpa normalisasi data. Model ini menghasilkan nilai rata rata Shillouete Score sebesar 0,8176. Dominasi nilai s(i) yang berada pada rentang tinggi , termasuk beberapa nilai mencapai nilai maximum 1 mengindikasi tingkat kohesi internal yang baik, dan jarak a(i) secara konsisten lebih kecil daripada jarak b(i) meskipun ada beberapa nilai s(i) yang rendah dan berada pada area transisi antar cluster. Proses perhitungan yang sama juga diterapkan pada dua metrik jarak dan dua skenario data. Berikut ditampilkan hasil klaster yang dihasilkan menggunakan K Medoids pada tabel 8**.**

Tabel 8. Hasil Cluster K-Medoids

2021				2022				2023			
Euclidean Tanpa Normalisasi	Euclidean Normalisasi	Manhattan Tanpa Normalisasi	Manhattan Normalisasi	Euclidean Tanpa Normalisasi	Euclidean Normalisasi	Manhattan Tanpa Normalisasi	Manhattan Normalisasi	Euclidean Tanpa Normalisasi	Euclidean Normalisasi	Manhattan Tanpa Normalisasi	Manhattan Normalisasi
2	2	1	2	1	1	1	1	2	1	2	2
2	2	1	2	1	1	1	1	1	2	1	1
2	2	1	2	1	1	1	1	1	2	1	1
2	2	1	2	1	1	1	1	2	2	2	1
2	2	2	1	2	1	1	2	2	2	2	1
2	2	1	2	1	1	1	1	1	2	1	1
2	2	1	2	1	1	1	1	1	2	1	1
2	2	2	1	1	1	1	1	2	1	2	2
2	2	1	2	1	1	1	1	1	2	1	1
2	2	2	1	2	2	2	2	2	1	2	2
2	2	1	2	1	1	1	1	1	2	1	1
2	2	1	2	1	1	1	1	1	2	1	1
1	1	2	2	2	1	2	1	2	2	2	1

JURIKOM (Jurnal Riset Komputer), Vol. 13 No. 1, Februari 2026
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v13i1.9520
Hal 451-461
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

2	2	1	2	1	1	1	1	1	2	1	1

Hasil pengelompokkan pada tabel 8 menggunakan metode K-Medoids pada rentang 2021-2023 menunjukkan pola klaster yang relatif stabil di semua tahun meskipun menggunakan metrik jarak berbeda. Tanpa normalisasi, Euclidean menghasilkan klaster yang lebih seragam, sedangkan Manhattan memperlihatkan variasi lebih besar karena lebih sensitif terhadap perbedaan nilai antar data. Konsistensi pola antar tahun mengindikasikan bahwa distribusi faktor penyebab perceraian tidak mengalami perubahan berarti, sementara perbedaan kecil yang muncul pada 2022 terutama pada metrik Manhattan menunjukkan adanya dinamika data pada tahun tersebut.

3.6 Hasil Evaluasi dan Komparasi Silhouette Score

Untuk mengevaluasi kualitas hasil klasterisasi, dilakukan perhitungan Silhouette Score pada seluruh metode yang digunakan, baik pada skenario tanpa normalisasi atau dengan normalisasi data. Hasil evaluasi Silhouette Score pada kedua skenario pengolahan data disajikan pada tabel berikut.

Tabel 9. Hasil Evaluasi dan Komparasi Silhouette Score

Metode Klasterisasi	Tanpa Normalisasi	Dengan Normalisasi
K-Means Euclidean	0,8344	0,834389
K-Means Manhattan	0,8293	0,849547
K-Medoids Euclidean	0,8176	0,572784
K-Medoids Manhattan	0,8293	0,624594

Berdasarkan tabel 9 diatas, hasil evaluasi menunjukkan bahwa pada skenario tanpa normalisasi seluruh metode klasterisasi menghasilkan rata-rata Silhouette Score yang melebihi 0,8 yang mengindikasikan kualitas kelompok yang baik, metode K-Means Manhattan dengan normalisasi menunjukkan performa terbaik dengan nilai rata-rata Silhouette sebesar 0,849547, Sementara itu, pada algoritma K-Medoids, konfigurasi yang memberikan hasil paling baik adalah jarak Manhattan tanpa normalisasi, dengan nilai rata-rata Silhouette sebesar 0,8293. Sedangkan nilai terendah diperoleh oleh K-Medoid Euclidean dengan normalisasi sebesar 0,572784.

3.7 Analisis Perpindahan Klaster

Analisis perpindahan klaster dilakukan dengan menggunakan metode clustering yang terbaik yaitu K-Means Manhattan dengan normalisasi Min-Max Scaler, yang telah ditetapkan berdasarkan nilai rata-rata Shilouette Score tertinggi pada tahap sebelumnya. Metode ini diterapkan secara konsisten pada setiap tahun pengamatan, yaitu 2021,2022,2023, kemudian hasil pengelompokkan antar tahun dibandingkan untuk mengidentifikasi perpindahan anggota klaster. Berikut hasil analisis klaster pada tabel 10.

Tabel 10. Analisis Klaster K-Means Manhattan Normalisasi

2021	2022	2023	faktor
1	1	2	zina
1	1	2	mabuk
1	1	2	madat
1	1	2	judi
1	1	2	meninggalkan satu pihak
1	1	2	dihukum penjara
1	1	2	poligami
1	1	2	KDRT
1	1	2	Cacat badan
2	2	1	perselisihan
1	1	2	kawin paksa
1	1	2	murtad
1	1	2	ekonomi
1	1	2	lain lain
1	1	2	zina

Pada tabel 10 diatas menunjukkan hasil analisis perpindahan klaster menggunakan K-Means Manhattan dengan normalisasi menunjukkan bahwa pengelompokan faktor pada periode 2021–2023 membentuk pola yang konsisten. Perbedaan label klaster yang muncul, khususnya pada tahun 2023, tidak mencerminkan perubahan struktur pengelompokan, melainkan hanya perbedaan penomoran klaster yang dapat terjadi dalam proses *clustering*. Dengan demikian, faktor-faktor yang dianalisis tetap berada dalam kelompok dengan karakteristik yang sama, sehingga tidak terdapat perubahan pola klaster yang signifikan antar tahun.

3.8 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma K-Means dan K-Medoids mampu membentuk klaster faktor penyebab. Hasil klasterisasi menunjukkan bahwa perbedaan metode dan metrik jarak sangat memengaruhi kualitas pemisahan data. K-Means dan K-Medoids merespons variasi nilai dengan cara yang berbeda, K-Means lebih peka terhadap perubahan skala sehingga kinerjanya meningkat setelah normalisasi, sementara K-Medoids lebih stabil pada data asli karena tidak bergantung pada nilai rata-rata, tetapi kualitas klasternya tidak setinggi K-Means. Perbedaan ini tampak jelas pada nilai Silhouette. K-Means Manhattan dengan normalisasi menghasilkan nilai tertinggi **(**0,849547**)** dan menjadi konfigurasi paling konsisten di seluruh skenario. Manhattan Distance bekerja lebih baik karena mampu menangkap perbedaan antar faktor yang tidak seragam, sedangkan Euclidean lebih mudah berubah ketika rentang data diubah melalui normalisasi. Pada K-Medoids, performa terbaik muncul pada Manhattan tanpa normalisasi **(**0,8293**)**, sementara nilai terendah muncul pada Euclidean dengan normalisasi **(**0,572784), menandakan bahwa kombinasi tersebut tidak cocok karena menurunkan ketegasan batas klaster.
Struktur klaster yang terbentuk juga memperlihatkan kecenderungan berbeda antar metode. Manhattan cenderung menjaga kestabilan klaster pada K-Means maupun K-Medoids, sedangkan Euclidean lebih sering menyebabkan perpindahan anggota klaster pada beberapa tahun pengamatan. Hal ini mengindikasikan bahwa pola kedekatan antar faktor lebih cocok direpresentasikan oleh Manhattan Distance.

Analisis perpindahan klaster menggunakan metode terbaik yaitu K-Means Manhattan dengan normalisasi yang menunjukkan bahwa pola klaster relatif stabil sepanjang 2021–2023. Perbedaan label pada 2023 hanya merupakan perbedaan penomoran klaster, bukan perubahan struktur kelompok. Ditinjau dari struktur klaster yang terbentuk melalui metode optimal tersebut, faktor perselisihan terus-menerus secara konsisten berada pada klaster dengan karakteristik paling menonjol. Posisi faktor ini menunjukkan bahwa dominansi suatu faktor dalam proses klasterisasi tidak ditentukan oleh besarnya nilai satu variabel, melainkan oleh kesamaan pola multidimensi antar faktor**.**

4. KESIMPULAN

Hasil analisis menunjukkan bahwa K-Means dengan jarak Manhattan dan normalisasi Min–Max merupakan metode yang paling optimal, ditunjukkan oleh nilai Silhouette Score tertinggi sebesar 0,8495. Metode ini menghasilkan klaster yang lebih terstruktur dan stabil dibandingkan menggunakan metode lainnya. Berdasarkan pengelompokan yang dihasilkan oleh metode tersebut, faktor perselisihan terus-menerus muncul sebagai faktor yang paling dominan karena struktur klaster yang terbentuk konsisten dari tahun ke tahun. Sebaliknya, faktor penyebab lainnya masih menunjukkan ketidakstabilan, karena posisinya dalam klaster sering berpindah atau mengalami perubahan komposisi antar tahun. Dengan demikian, K-Means dengan jarak Manhattan dan normalisasi Min–Max dapat dinyatakan sebagai metode klasterisasi yang paling optimal dalam menganalisis faktor penyebab perceraian pada penelitian ini. Dalam penelitian selanjutnya disarankan untuk memperluas variabel dan membandingkan dengan pendekatan metode *clustering* yang lain guna memperoleh hasil yang lebih komprehensif.

REFERENCES

[1] S. Sofiatun, F. Deviantony, and W. Kusumawardani, "A Systematic Review of the Psychological and Social Consequences of Parental Divorce on Children," *J. Ilmu-Ilmu Kesehat.*, vol. 11, no. 1, pp. 6–13, 2025, doi: 10.52741/jiikes.v11i1.120.
[2] M. P. A. Tubagus Fajar Setiadi,Lukman Aqil Prambudi, "Disintegrasi Rumah Tangga : Faktor Perceraian dalam Tinjauan," *Marital J. Huk. Kel. Islam*, vol. 3, no. 2, 2025, doi: 10.35905/marital.
[3] P. Agama and K. Kebumen, "Article Info Accepted Manusia sebagai makhluk sosial memiliki kodrat untuk berinteraksi dan bersosialisasi dengan sesamanya . Dalam pengertian lain , untuk menjamin kelangsungan hidupnya manusia membutuhkan bantuan dan perlu bekerja sama dengan manusia la," pp. 1–11, 2024.
[4] Y. Yendrizal and Y. Alfero, "Data Mining Penjualan Bibit Tanaman Menggunakan Algoritma Apriori," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 7, no. 6, pp. 1908–1917, 2024, doi: 10.32672/jnkti.v7i6.8346.
[5] D. Anggraini and M. S. Hasibuan, "Initial Centroid Pada Algoritma K-Means Dan K-Medoids," *Jatilima J. Multimed. Dan Teknol. Inf.* , vol. 7, no. 3, pp. 487–500, 2025.
[6] G. Gan, "A k-Means Algorithm with Automatic Outlier Detection," *Electron.*, vol. 14, no. 9, pp. 1–17, 2025, doi: 10.3390/electronics14091723.
[7] M. Orisa and A. Faisol, "Analisis Algoritma Partitioning Around Medoid untuk Penentuan Klasterisasi," *J. Teknol. Inf. dan Terap.*, vol. 8, no. 2, pp. 86–90, 2021, doi: 10.25047/jtit.v8i2.258.
[8] M. A. C. Roring, P. A. Ginoga, R. Ibrahim, R. C. Rompas, A. Yusupa, and S. D. E. Paturusi, "E-issn : 2988-1986," *Multidisiplin Saintek*, vol. 8, no. 5, pp. 1–16, 2025.
[9] S. Komputer and E. Yanti, "Analisis Algoritma K-Means Dalam Pengelompokan Perkara Perceraian Berdasarkan Kelurahan Di Kota Jambi," vol. 16, no. 1, pp. 9–19, 2021.
[10] H. D. Tampubolon, S. Suhada, M. Safii, S. Solikhun, and D. Suhendro, "Penerapan Algoritma K-Means dan K-Medoids Clustering untuk Mengelompokkan Tindak Kriminalitas Berdasarkan Provinsi," *J. Ilmu Komput. dan Teknol.*, vol. 2, no. 2, pp. 6–12, 2021, doi: 10.35960/ikomti.v2i2.703.
[11] F. Zahro, "Penentuan klaster terbaik dengan elbow method dalam faktor penyebab perceraian menggunakan k-means skripsi," 2024.
[12] A. Triayudi, Iksal, and R. Haerani, "Data Mining K-Means Algorithm for Performance Analysis," *J. Phys. Conf. Ser.*, vol. 2394, no. 1, pp. 0–7, 2022, doi: 10.1088/1742-6596/2394/1/012031.
[13] A. Druzhinin, R. W. Hobbs, and X.-Y. Li, "The Impact of Data Pre-Processing on Sub-Basalt Seismic Imaging," vol. 15, no. August, pp. 540–544, 2022, doi: 10.3997/2214-4609-pdb.3.p305.
[14] D. A. Nasution, H. H. Khotimah, and N. Chamidah, "Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *Comput. Eng. Sci. Syst. J.*, vol. 4, no. 1, p. 78, 2019, doi: 10.24114/cess.v4i1.11458.
[15] R. Sammouda and A. El-Zaart, "An Optimized Approach for Prostate Image Segmentation Using K-Means Clustering Algorithm

with Elbow Method," *Comput. Intell. Neurosci.*, vol. 2021, 2021, doi: 10.1155/2021/4553832.
[16] D. Xu and Y. Tian, "A Comprehensive Survey of Clustering Algorithms," *Ann. Data Sci.*, vol. 2, no. 2, pp. 165–193, 2015, doi: 10.1007/s40745-015-0040-1.
[17] R. D. Rudianto and A. W. Wijayanto, "Analisis Perbandingan K-Means dan K-Medoids dalam Pengelompokan Provinsi Berdasarkan Indeks Demokrasi Indonesia 2021," *Komputika J. Sist. Komput.*, vol. 13, no. 1, pp. 19–26, 2024, doi: 10.34010/komputika.v13i1.10812.
[18] Y. Sari, A. R. Baskara, and P. B. Prakoso, "Penerapan Metode K-Means Berbasis Jarak untuk Deteksi Kendaraan Bergerak," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 4, pp. 683–690, 2022, doi: 10.25126/jtiik.2022945768.
[19] T. Rahmawati, Y. Wilandari, and P. Kartikasari, "Analisis Perbandingan Silhouette Coefficient Dan Metode Elbow Pada Pengelompokkan Provinsi Di Indonesia Berdasarkan Indikator Ipm Dengan K-Medoids," *J. Gaussian*, vol. 13, no. 1, pp. 13–24, 2024, doi: 10.14710/j.gauss.13.1.13-24.
[20] F. Irawati, A. P. W. Nugroho, and A. Wibowo, "Analisis Ekspor Kopi Menggunakan Clustering K-Means dan Davies-Bouldin Index," *J. Ilm. FIFO*, vol. 17, no. 2, p. 164, 2025, doi: 10.22441/fifo.2025.v17i2.006.
[21] B. Hartono, S. Eniyati, and K. Hadiono, "Perbandingan Metode Perhitungan Jarak pada Nilai Centroid dan Pengelompokan Data Menggunakan K-Means Clustering," *J. Sist. Komput. dan Inform.*, vol. 4, no. 3, p. 503, 2023, doi: 10.30865/json.v4i3.6021.
[22] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. I. R.H.Zer, and D. Hartama, "Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia," *J. Teknol. Inf.*, vol. 4, no. 1, pp. 166–173, 2020, doi: 10.36294/jurti.v4i1.1296.
[23] R. Meiyanti, M. M. Munauwar, R. Fitria, and H. Al Kautsar, "Implementasi Algoritma K-Medoid pada Clustering Sayuran Unggulan di Kabupaten Aceh Utara," *Teknika*, vol. 19, no. x, pp. 327–337, 2024.
[24] D. D. Abdurrahman, F. Agus, and G. M. Putra, "Implementasi Algoritma Partitioning Around Medoids (PAM) untuk Mengelompokkan Hasil Produksi Komoditi Perkebunan (Studi Kasus: Dinas Perkebunan Provinsi Kalimantan Timur)," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 16, no. 2, p. 130, 2021, doi: 10.30872/jim.v16i2.6520.
[25] B. N. Yulisasih, H. Herman, S. Sunardi, and H. Yuliansyah, "Evaluation of K-Means Clustering Using Silhouette Score Method on Customer Segmentation," *Ilk. J. Ilm.*, vol. 16, no. 3, pp. 330–342, 2024, doi: 10.33096/ilkom.v16i3.2325.330-342.