MATHunesa
Jurnal Ilmiah Matematika *Volume 13 No 02*
e-ISSN : 2716-506X | p-ISSN : 2301-9115 *Tahun 2025*

ANALISIS PERBANDINGAN METODE K-MEDOID DAN WARD DALAM KLASTERISASI NASABAH BANK CHURNERS

Etis Sunandi
Program Studi Statistika, Fakultas MIPA, Universitas Bengkulu, Kota Bengkulu, Indonesia
e-mail: esunandi@unib.ac.id*

Aditama Rouliber Simanullang
Program Studi Statistika, Fakultas MIPA, Universitas Bengkulu, Kota Bengkulu, Indonesia

Vivi Elvira Saputri Syah
Program Studi Statistika, Fakultas MIPA, Universitas Bengkulu, Kota Bengkulu, Indonesia

Della Nur Afni
Program Studi Statistika, Fakultas MIPA, Universitas Bengkulu, Kota Bengkulu, Indonesia

Resi Popita
Program Studi Statistika, Fakultas MIPA, Universitas Bengkulu, Kota Bengkulu, Indonesia

Winalia Agwil
Program Studi Statistika, Fakultas MIPA, Universitas Bengkulu, Kota Bengkulu, Indonesia

Abstrak

Penelitian ini membandingkan metode klasterisasi *K-Medoid* dan *Ward* dalam mengelompokkan nasabah Bank Churners. Data yang digunakan berasal dari *dataset* Bank Churners yang terdiri dari 5.000 pengamatan dan enam variabel terkait penggunaan kartu kredit. Hasil penelitian menunjukkan bahwa metode *Ward* memiliki koefisien *silhouette* lebih tinggi dibandingkan *K-Medoid*, sehingga lebih efektif dalam membentuk klaster yang homogen. *Ward* menghasilkan empat klaster dengan karakteristik berbeda, yang dapat membantu bank dalam menyusun strategi retensi nasabah. Berdasarkan hasil analisis, disarankan strategi promosi dan personalisasi layanan sesuai dengan karakteristik masing-masing klaster untuk mengurangi tingkat *churn*.
Kata Kunci: Klasterisasi, *K-Medoid*, *Ward*.

Abstract

This study compares the K-Medoid and Ward clustering methods in segmenting bank churner customers. The dataset used is the Bank Churners dataset, consisting of 5.000 observations and six variables related to credit card usage. The results show that the Ward method has a higher silhouette coefficient compared to K-Medoid, making it more effective in forming homogeneous clusters. Ward clustering produces four distinct customer segments, which can help banks develop targeted retention strategies. Based on the analysis, promotional strategies and personalized services are recommended to reduce customer churn according to each cluster's characteristics.
Keywords: *Clustering, K-Medoid, Ward.*

PENDAHULUAN

Selama lima belas tahun terakhir, jumlah nasabah bank terus mengalami peningkatan dari waktu ke waktu. Hal ini mendorong bank untuk tetap mempertahankan kualitas layanan yang mereka berikan. Saat ini, industri perbankan menghadapi perubahan besar dan kompleks, serta beragam tantangan. Kemajuan teknologi memungkinkan akses informasi menjadi lebih mudah, terutama terkait produk dan layanan perbankan. Akibatnya, nasabah lebih mudah berpindah ke bank pesaing (Nugraha and Syarif 2023).

Perpindahan nasabah ke bank lain dikenal sebagai *Churn*, yaitu ketika nasabah memutuskan untuk berhenti menggunakan layanan suatu bank dan beralih ke bank lain yang dianggap menawarkan layanan lebih baik. Fenomena *Customer Churn* perlu segera ditangani untuk mencegah dampak negatif bagi bank. Semakin tinggi tingkat *churn*, semakin penting bagi bank untuk mengevaluasi kualitas layanan yang diberikan. Hal ini dikarenakan menarik nasabah baru umumnya membutuhkan biaya lebih besar dibandingkan mempertahankan nasabah yang sudah ada (Miryam Clementine and Arum 2022).

Analisis klaster terbagi menjadi dua metode utama, yaitu *hierarchical method* dan *non-hierarchical method*. Pada *hierarchical method*, jumlah klaster yang dihasilkan belum ditentukan sejak awal, sementara dalam *non-hierarchical method*, jumlah klaster diasumsikan terlebih dahulu sebanyak *k* kelompok. Metode *hierarchical* mencakup beberapa teknik, seperti *Single Linkage, Complete Linkage, Centroid Linkage, Average Linkage,* dan *Ward's Method*. Sementara itu, metode yang termasuk dalam *non-hierarchical method* antara lain *K-Means* dan *K-Medoids* (Raja, Tinungki, and Sirajang 2024).

Dalam penelitian ini, digunakan metode *non-hierarchical* yaitu *K-Medoids*. *K-Medoids* adalah teknik klasterisasi berbasis partisi yang mengelompokkan sekumpulan objek ke dalam beberapa klaster. Metode ini menggunakan *medoid* sebagai pusat klaster, di mana objek-objek yang berada di sekitarnya dikelompokkan untuk membentuk klaster baru. Salah satu keunggulan *K-Medoids* adalah kemampuannya mengatasi kelemahan *K-Means*, yang rentan terhadap *outlier*. Selain itu, hasil klasterisasi *K-Medoids* tetap konsisten meskipun urutan data diacak (Zahra, Khalif, and Sari 2024).

Selain itu, untuk memperoleh perbandingan dengan metode lain, penelitian ini juga menggunakan *hierarchical method*, yaitu metode *Ward*. Metode *Ward* merupakan teknik klasterisasi yang populer karena berfokus pada minimisasi *varians* dalam klaster serta meminimalkan *Sum of Squares Error* (SSE). Keunggulan utama metode ini adalah kemampuannya dalam menghasilkan klaster yang lebih homogen dan terstruktur, sehingga efektif untuk berbagai aplikasi dalam analisis data (Andyani, Shobri, and Baihaqi 2024).

Dengan semakin meningkatnya persaingan di industri perbankan, memahami perilaku nasabah yang berpotensi melakukan *churn* menjadi sangat penting. Analisis klasterisasi dapat membantu bank dalam mengidentifikasi pola dan karakteristik nasabah yang cenderung berpindah layanan. Maka, dalam penelitian ini akan dibandingkan dua metode klasterisasi, yaitu *K-Medoid* sebagai metode *non-hierarchical* dan *Ward's Method* sebagai metode *hierarchical*, untuk menentukan pendekatan yang lebih efektif dalam klasterisasi nasabah Bank Churners.

KAJIAN TEORI

PRE-PROCESSING DATA

Pre-processing data adalah tahap penting dalam analisis data mining yang bertujuan untuk membersihkan, menyesuaikan format, serta mempersiapkan data agar lebih terstruktur dan akurat untuk proses analisis (Agung et al. 2023). *Pre-processing* data sering digunakan untuk mengurangi kesalahan serta bias sistematis dalam data mentah sebelum dilakukan analisis. Proses ini mencakup beberapa tahap, seperti seleksi data, pembersihan (*cleaning*), normalisasi, transformasi, dan pelatihan data (*training data*) (Satriatama et al. 2023).

NORMALISASI DATA

Metode normalisasi data adalah suatu proses yang bertujuan untuk menyamakan rentang nilai beberapa variabel, sehingga tidak ada nilai yang terlalu besar atau terlalu kecil. Hal ini akan memudahkan dalam melakukan analisis statistik (Kusnaidi, Gulo, and Aripin 2022). Salah satu langkah penting dalam pra-pemrosesan data adalah normalisasi. Tujuan dari proses normalisasi ini adalah untuk mengatur nilai-nilai dalam *dataset* agar berada dalam rentang yang sama (Ahmad Harmain et al. 2022).

DATA MINING

Data Mining adalah metode yang digunakan untuk menganalisis pola dan karakteristik guna memprediksi tren di masa depan serta mengungkap informasi tersembunyi yang sebelumnya tidak terlihat dalam kumpulan data yang besar. Proses ini mengeksplorasi pengetahuan dan pola dalam data menggunakan pendekatan statistik, matematika, serta machine learning. Dalam data mining, *clustering* merupakan salah satu teknik yang digunakan untuk menganalisis dan mengelompokkan data (Sekar

Setyaningtyas, Indarmawan Nugroho, and Arif 2022).

ANALISIS KLASTER

Analisis klaster merupakan salah satu alat penting dalam pengolahan data statistik yang digunakan untuk menganalisis dan mengelompokkan data. Metode ini secara otomatis mengelompokkan objek ke dalam klaster berdasarkan tingkat kemiripannya. Klaster yang ideal memiliki tingkat homogenitas yang tinggi di dalamnya serta heterogenitas yang tinggi dibandingkan dengan klaster lainnya (Dani, Wahyuningsih, and Rizki 2020).

METODE *K-MEDOID*

K-Medoids Clustering, atau yang juga dikenal sebagai *Partitioning Around Medoids* (PAM), merupakan varian dari metode *K-Means*. Metode ini menggunakan *medoid* sebagai pusat klaster, bukan rata-rata seperti pada *K-Means*, sehingga hasil klasterisasi menjadi lebih tahan terhadap nilai ekstrem dalam kumpulan data. Pendekatan ini bertujuan untuk menghasilkan partisi yang lebih stabil dan kurang sensitif terhadap *outlier* (Rohman and Wibowo 2024). Algoritma *K-Medoids* merupakan metode *partitional clustering* yang membagi sekumpulan N objek ke dalam K klaster. Dalam algoritma ini, klaster diwakili oleh objek yang dipilih dari kumpulan data, yang disebut *medoid*. Pembentukan klaster dilakukan dengan mengukur kedekatan antara *medoid* dan objek *non-medoid*, sehingga setiap objek dikelompokkan ke dalam klaster dengan *medoid* terdekatnya (Kholifah, Bahri, and Matematika 2025). Tahapan penyelesaian *K-Medoids* adalah sebagai berikut (Sulistyawati and Sadikin 2021):

a. Menginisialisasi pusat klaster sebanyak jumlah klaster (k).
b. Mengalokasikan setiap objek ke klaster terdekat menggunakan *Euclidean Distance*, dengan rumus:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \qquad (1)$$

c. Memilih objek secara acak dalam setiap klaster sebagai calon *medoid* baru.
d. Menghitung jarak setiap objek dalam klaster terhadap calon *medoid* baru.
e. Menghitung total simpangan (S), yaitu selisih antara total jarak baru dengan total jarak sebelumnya. Jika $S < 0$, maka dilakukan pergantian *medoid* dengan objek dalam klaster yang baru terbentuk.
f. Mengulangi langkah c hingga e sampai tidak ada perubahan *medoid*. Jika tidak ada perubahan, maka klaster akhir dan anggotanya telah diperoleh.

METODE *WARD*

Metode *Ward* merupakan teknik pengelompokan yang menggunakan pendekatan analisis *varians* untuk menentukan jarak antar klaster dengan meminimalkan jumlah kuadrat dalam proses perhitungan (Imasdiani, Purnamasari, and Amijaya 2022). Proses pengelompokan dalam metode *Ward* didasarkan pada kriteria *Sum of Squares Error* (SSE), yang mengukur tingkat kehomogenan antara dua objek dengan meminimalkan jumlah kuadrat kesalahan. Nilai SSE untuk klaster yang hanya memiliki satu objek atau satu item bernilai nol. Jika terdapat N item yang dikelompokkan dalam satu klaster, maka nilai SSE dapat dihitung menggunakan persamaan berikut (Irwan, Sanusi, and Hasanah 2024):

$$SSE = \sum_{i=1}^{n}(X_i - \bar{X})'\,(X_i - \bar{X}) \qquad (2)$$

di mana:

X_i = nilai data ke-i dengan $i = 1,2,3,\ldots,n$.
$\bar{X}$ = vektor kolom yang berisi rata-rata nilai pengamatan dalam klaster.
n = jumlah total pengamatan dalam klaster.

METODE

DATA PENELITIAN

Penelitian ini menggunakan *dataset* Bank Churners, yang berisi informasi tentang nasabah sebuah bank terkait penggunaan kartu kredit mereka. Data ini diperoleh dari Kaggle *https://www.kaggle.com/datasets/imanemag/bankchurnerscsv*. Terdapat total 5.000 pengamatan dan 6 Variabel yang digunakan. Di mana, variabel-variabel yang akan digunakan dalam penelitian ini sebagai berikut:

Tabel 1. Variabel Penelitian

Variabel	Satuan
Clientnum	ID
Customer Age	Tahun
Credit Limit	USD
Total Trans Amt	USD

Total Trans Ct	Jumlah Transaksi
Avg Utilization Ratio	Persentase (%)

RANCANGAN PENELITIAN

Rancangan penelitian digunakan untuk mencapai tujuan dari penelitian yang dilakukan. Langkah langkah yang dilakukan dalam perbandingan metode *K-Medoid* dan *Ward* dengan bantuan *software* R adalah sebagai berikut:

1. Mengumpulkan dan menyiapkan data
   Data nasabah Bank Churners.
2. *Pre-Processing* dan Normalisasi data
   Data dinormalisasi agar skala antar variabel seimbang dan tidak mendominasi hasil klasterisasi.
3. Penerapan metode *K-Medoid*
   Langkah-langkah yang harus dilakukan untuk metode *K-Medoid* adalah sebagai berikut:
   a. Tentukan jumlah klaster (k).
   b. Hitung jarak setiap objek ke semua medoid menggunakan Jarak *Euclidean*.
   c. Pemilihan *medoid* baru (Pertukaran).
   d. Evaluasi *medoid* baru.
   e. Perhitungan total Simpangan (S).
   f. Ulangi sampai tidak ada perubahan medoid.
4. Penerapan metode *Ward*
   Langkah-langkah yang harus dilakukan untuk metode *Ward* adalah sebagai berikut:
   a. Setiap data dianggap sebagai klaster tunggal.
   b. Hitung SSE antara setiap pasangan klaster.
   c. Gabungkan dua klaster dengan kenaikan SSE terkecil.
   d. Ulangi proses hingga jumlah klaster sesuai target.
   e. Hasil divisualisasikan dalam bentuk dendrogram.
5. Evaluasi hasil klasterisasi
6. Interpretasi hasil

HASIL DAN PEMBAHASAN

HASIL *CLUSTERING*

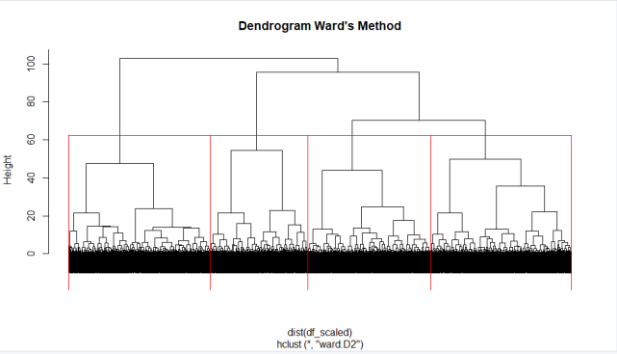
Analisis *clustering* menggunakan algoritma *K-Medoid* dan *Ward* menggunakan data yang sudah dinormalisasikan terlebih dahulu agar mendapatkan hasil yang optimal. Pada analisis *clustering* menggunakan algoritma *K-Medoid* dan *Ward* dengan bantuan *software* R, langkah pertama yang dilakukan adalah menentukan jumlah *cluster*. Dalam penelitian ini digunakan metode uji akurasi data menggunakan *silhouette score* untuk mengidentifikasi jumlah *cluster* terbaik pada algoritma *K-Medoid* dan *Ward*. Selain itu dilakukan pula pemilihan model terbaik dengan jumlah *cluster* masing-masing tiap metode dengan menggunakan koefisisen *silhouette* terbaik. Hasil *clustering* akan semakin baik jika nilai koefisien *silhouette* yang dihasilkan mendekati 1.

Tabel 2. Perbandingan Hasil *Clustering*

Clustering	**Jumlah *Cluster***	***Silhouette***
K-Medoid	4	0,694
Ward	4	0,712

Langkah selanjutnya dalam algoritma *Ward* adalah mementukan nilai *centroid* atau pusat *cluster* dari data yang didapatkan dengan jumlah *cluster* yang telah ditentukan yaitu 4 *cluster*. Diperoleh hasil *clustering* metode *Ward* menunjukkan empat karakteristik kelompok pelanggan yang berbeda.

1. Pada *cluster* 1 terdiri dari pelanggan dengan rata-rata usia paling muda 37 tahun, limit kredit sedang, dan jumlah transaksi yang relatif rendah, namun memiliki rasio pemanfaatan kartu yang cukup tinggi, menandakan penggunaan kartu yang cukup aktif sesuai limitnya.
2. *Cluster* 2 berisi pelanggan dengan usia rata-rata 47 tahun, memiliki limit kredit paling besar (24.946,67), tetapi dengan rasio pemanfaatan yang sangat rendah (0,05), menggambarkan kelompok pelanggan yang diberikan kepercayaan tinggi tetapi cenderung jarang memaksimalkan penggunaannya.
3. *Cluster* C3 diisi oleh pelanggan dengan usia rata-rata 54 tahun, memiliki limit kredit dan jumlah transaksi terendah, tetapi pemanfaatan kartunya cukup tinggi, menandakan penggunaan yang konsisten meskipun dalam jumlah kecil.
4. Terakhir, *cluster* C4 berisi pelanggan berusia menengah 47 tahun dengan limit kredit kecil, namun memiliki jumlah dan frekuensi transaksi tertinggi, serta rasio pemanfaatan kartu yang tinggi, menunjukkan kelompok pelanggan yang sangat aktif dalam menggunakan kartu kredit meskipun dengan batasan limit yang lebih kecil.



Gambar 1. Dendogram *Ward Clustering*

Gambar di atas menunjukkan visualisasi hasil analisis menggunakan *Ward clustering*. Berdasarkan hasil tersebut dapat diketahui hasil pengelompokkan pada tabel berikut.

Tabel 3. Hasil *Ward Clustering*

Cluster	**Total**
C1	1397
C2	974
C3	1223
C4	1406

Berdasarkan Tabel 3, hasil clustering menggunakan metode Ward menghasilkan 4 kelompok (*cluster*) dengan jumlah anggota yang berbeda-beda.

a. *Cluster* 1 terdiri dari 1.397 nasabah, dengan contoh *client* seperti 768805383 dan 818770208, menunjukkan kelompok yang cukup besar.
b. *Cluster* C2 memiliki jumlah anggota paling sedikit, yaitu 974 nasabah, menandakan kelompok dengan karakteristik tertentu yang hanya dimiliki oleh sebagian kecil pelanggan.
c. *Cluster* C3 memiliki 1.223 anggota dengan contoh *clientnum* 710986133 dan 759952688, menunjukkan kelompok yang cukup besar tetapi tidak sebesar C1 dan C4.
d. Sementara itu, *cluster* C4 menjadi kelompok dengan jumlah anggota terbanyak, yaitu 1.406 nasabah, menunjukkan bahwa karakteristik pelanggan dalam *cluster* ini paling dominan di data.

UCAPAN TERIMA KASIH

Kami mengucapkan terima kasih kepada semua pihak yang telah berkontribusi dalam penelitian ini. Dukungan dari keluarga, dosen pembimbing, sahabat, serta berbagai pihak yang telah memberikan masukan dan saran sangat membantu dalam penyusunan artikel ini.

PENUTUP

SIMPULAN

Berdasarkan penelitian yang telah dilakukan, metode *Ward* lebih baik ketimbang *K-Medoid* karena memiliki Koefisien *Silhouette* lebih besar (mendekati 1) yang berjumlah 4 *cluster*. Solusi untuk setiap *cluster* adalah untuk *cluster* 1, disarankan memberikan promosi menarik dan program *reward* untuk mendorong peningkatan transaksi. *Cluster* 2 memerlukan pendekatan personal melalui penawaran eksklusif dan fasilitas premium agar lebih aktif menggunakan kartu. Pada *cluster* 3, perusahaan dapat menawarkan peningkatan limit kredit dan layanan loyalitas untuk mempertahankan penggunaan konsisten. Sementara itu, *cluster* 4 sebaiknya diberikan kenaikan limit kredit secara bertahap dan promo khusus, mengingat nasabah sudah sangat aktif menggunakan kartu meskipun dengan limit kecil.

SARAN

Saran untuk penelitian selanjutnya dapat ditambahkan variabel yang mempengaruhi dengan metode pengelompokkan yang berbeda. Sehingga tidak menutup kemungkinan hasil yang didapatkan juga akan berbeda dari penelitian ini.

DAFTAR PUSTAKA

Agung, Anak, Aryasatya Daniswara, I. Kadek, and Dwi Nuryana. 2023. "Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru." *Journal of Informatics and Computer Science* 05:97–100.

Ahmad Harmain, Paiman Paiman, Henri Kurniawan, Kusrini Kusrini, and Dina Maulina. 2022. "Normalisasi Data Untuk Efisiensi K-Means Pada Pengelompokan Wilayah Berpotensi Kebakaran Hutan Dan Lahan Berdasarkan Sebaran Titik Panas." *TEKNIMEDIA: Teknologi Informasi Dan Multimedia* 2(2):83–89. doi: 10.46764/teknimedia.v2i2.49.

Andyani, Rhavida Anniza, Muhammad Qolbi Shobri, and Muhamad Adzib Baihaqi. 2024. "Aplikasi Metode Ward Dengan Berbagai Pengukuran Jarak (Studi Kasus : Klasifikasi Tingkat Perekonomian Di Indonesia)." (4):177–90. doi: 10.56972/jikm.v4i2.208.

Dani, Andrea Tri Rian, Sri Wahyuningsih, and Nanda

Arista Rizki. 2020. “Pengelompokkan Data Runtun Waktu Menggunakan Analisis Cluster (Studi Kasus: Nilai Ekspor Komoditi Migas Dan Nonmigas Provinsi Kalimantan Timur Periode Januari 2000-Desember 2016).” *Jurnal EKSPONENSIAL* 11(1):29–38.

Imasdiani, Imasdiani, Ika Purnamasari, and Fidia Deny Tisna Amijaya. 2022. “Perbandingan Hasil Analisis Cluster Dengan Menggunakan Metode Average Linkage Dan Metode Ward.” *Eksponensial* 13(1):9. doi: 10.30872/eksponensial.v13i1.875.

Irwan, Irwan, Wahidah Sanusi, and Afifatun Hasanah. 2024. “Perbandingan Analisis Cluster Metode Complete Linkage Dan Metode Ward Dalam Pengelompokkan Indeks Pembangunan Manusia Di Sulawesi Selatan.” *Journal of Mathematics, Computations and Statistics* 7(1):75–86. doi: 10.35580/jmathcos.v7i1.2089.

Kholifah, Fitri Nur, Saiful Bahri, and Prodi Matematika. 2025. “Pengelompokan Daerah Produksi Tanaman Biofarmaka Menurut Jenis Tanaman Dengan Metode K – Means Clustering.” 6(1):61–69. doi: 10.20956/ejsa.v6i1.27200.

Kusnaidi, Muhammad Rafli, Timotius Gulo, and Soeb Aripin. 2022. “Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal.” *Journal of Computer System and Informatics (JoSYC)* 3(4):330–38. doi: 10.47065/josyc.v3i4.2112.

Miryam Clementine, and Arum. 2022. “Prediksi Churn Nasabah Bank Menggunakan Klasifikasi NaÃ¯ve Bayes Dan ID3.” *Jurnal Processor* 17(1):9–18. doi: 10.33998/processor.2022.17.1.1170.

Nugraha, Wahyu, and Muhamad Syarif. 2023. “Teknik Weighting Untuk Mengatasi Ketidakseimbangan Kelas Pada Prediksi Churn Menggunakan XGBoost, LightGBM, Dan CatBoost.” *Techno.Com* 22(1):97–108. doi: 10.33633/tc.v22i1.7191.

Raja, Nur Alfianingsih, Georgina Maria Tinungki, and Nasrah Sirajang. 2024. “Implementasi Algoritma Centroid Linkage Dan K-Medoids Dalam Mengelompokkan Kabupaten/Kota Di Sulawesi Selatan Berdasarkan Indikator Pendidikan.” *ESTIMASI: Journal of Statistics and Its Application* 5(1):61–74. doi: 10.20956/ejsa.v5i1.13605.

Rohman, Nur, and Arief Wibowo. 2024. “Perbandingan Metode K-Medoids Dan Metode K-Means Dalam Analisis Segmentasi Pelanggan Mall.” *SINTECH (Science and Information Technology) Journal* 7(1):49–58. doi: 10.31598/sintechjournal.v7i1.1507.

Satriatama, Ananda Elang, Ari Prasetyo Wibowo, I. Gusti Ngurah Arnold, Reyhan Bayu Pratama, Tegar Alwinata Masyhuda, Yohannes Alexander Agusti, Endah Purwanti, and Indah Werdiningsih. 2023. “Analisis Klaster Data Pasien Diabetes Untuk Identifikasi Pola Dan Karakteristik Pasien.” *Jurnal Teknologi Dan Sistem Informasi Bisnis* 5(3):172–82. doi: 10.47233/jteksis.v5i3.828.

Sekar Setyaningtyas, Bangkit Indarmawan Nugroho, and Zaenul Arif. 2022. “Tinjauan Pustaka Sistematis: Penerapan Data Mining Teknik Clustering Algoritma K-Means.” *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang* 10(2):52–61. doi: 10.21063/jtif.2022.v10.2.52-61.

Sulistyawati, Anggi Ayu Dwi, and Mujiono Sadikin. 2021. “SISTEMASI: Jurnal Sistem Informasi Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan.” *SISTEMASI: Jurnal Sistem Informasi* 10(3):516–26.

Zahra, Fathimatuz, Assyifa Khalif, and Betha Nurina Sari. 2024. “Pengelompokan Tingkat Kemiskinan Di Setiap Provinsi Di Indonesia Menggunakan Algoritma K-Medoids.” *Jurnal Informatika Dan Teknik Elektro Terapan* 12(2):2830–7062.