Implementasi Data Mining Menggunakan Teknik Clustering dengan Metode K-Means

Fadhlulrahman Khalish[1], Nurul Maharani Piranti[2], Okky Martadireja[3]
[1,2,3]Politeknik Imigrasi, Indonesia
E-mail: frk.khalish@gmail.com, nurulpiranti@gmail.com, okkypm@gmail.com

Article Info
Article History
Received: 2025-03-11
Revised: 2025-04-27
Published: 2025-05-10

Keywords:
Data Mining;
Clustering;
K-Means.

Abstract
Data collection in data mining is a crucial stage in the Knowledge Discovery in Databases (KDD) process. Clustering techniques, particularly the K-Means algorithm, are often utilized to group data with similar characteristics. This literature review aims to analyze various approaches and implementations of the K-Means algorithm in data mining. The methods employed include a systematic study of relevant scientific publications, encompassing journals, conferences, and research articles. The review results indicate that the K-Means algorithm is effective in identifying patterns and structures within data but is sensitive to the initialization of cluster centers and the presence of outliers. Several studies propose modifications and optimizations of K-Means to address these limitations. In conclusion, K-Means remains a relevant clustering algorithm in data mining, with the potential for further development to enhance accuracy and efficiency.

Artikel Info
Sejarah Artikel
Diterima: 2025-03-11
Direvisi: 2025-04-27
Dipublikasi: 2025-05-10

Kata kunci:
Data Mining;
Clustering;
K-Means.

Abstrak
Pengumpulan data *mining* merupakan tahapan yang penting dalam proses *knowledge discovery in databases* (KDD). Teknik *clustering*, khususnya algoritma K-Means, sering dimanfaatkan untuk mengelompokkan data yang memiliki karakteristik yang sama. Tinjauan literatur ini bertujuan untuk menganalisis berbagai pendekatan dan implementasi algoritma K-Means dalam pengumpulan data *mining*. Metode yang digunakan diantara lain adalah studi sistematis terhadap publikasi ilmiah terkait, yang mencakup jurnal, konferensi, dan artikel penelitian. Hasil tinjauan menunjukkan bahwa algoritma K-Means efektif dalam mengidentifikasi pola dan struktur yang ada dalam data, namun sensitif terhadap inisialisasi pusat *cluster* dan keberadaan *outlier*. Beberapa penelitian mengusulkan modifikasi dan optimasi K-Means untuk mengatasi keterbatasan ini. Kesimpulannya, K-Means tetap menjadi algoritma *clustering* yang relevan dalam pengumpulan data *mining*, dengan potensi pengembangan lebih lanjut untuk meningkatkan akurasi dan efisiensi.

I. PENDAHULUAN

Dalam era digital ini, volume data yang dihasilkan setiap harinya semakin besar dan kompleks, mencakup berbagai format dan sumber, mulai dari transaksi bisnis, media sosial, hingga sensor IoT. Dalam menghadapi tantangan ini, *data mining* dapat menjadi solusi utama dalam menggali dan mengekstrak informasi berharga serta mengungkap pola tersembunyi yang dapat memberikan wawasan strategis bagi pengambilan keputusan. Proses *Knowledge Discovery in Databases* (KDD) terdiri atas serangkaian tahapan yang sistematis, dimulai dari pengumpulan data, pembersihan, transformasi, sampai analisis dan interpretasi hasil. Di antaranya pengumpulan data merupakan langkah awal yang sangat krusial, karena kualitas dan kelengkapan data yang diperoleh akan menentukan validitas serta efektivitas hasil analisis yang dilakukan dalam penelitian maupun penerapan praktisnya di berbagai bidang.

Teknik *clustering* merupakan salah satu metode yang utama dalam *data mining*, yang mana berfungsi untuk mengelompokkan data ke dalam beberapa *cluster* berdasarkan tingkat kemiripan karakteristik yang dimilikinya. Dengan pendekatan ini, data yang memiliki karakteristik sama akan dikelompokkan ke dalam satu *cluster*, namun data dengan karakteristik berbeda akan ditempatkan dalam *cluster* yang berbeda. Salah satu algoritma *clustering* yang paling populer dan banyak digunakan adalah *K-Means*. Algoritma ini bekerja dengan membagi data ke dalam sejumlah *cluster* yang telah ditentukan sebelumnya, dengan tujuan mengurangi jarak antara data dalam satu *cluster* terhadap pusatnya (*centroid*). Karena efisiensinya dalam menangani data dalam jumlah yang besar dan kemampuannya menghasilkan pengelompokan yang akurat, *K-Means* dapat

diterapkan dalam berbagai bidang, seperti segmentasi pelanggan dalam bisnis dan pemasaran, analisis citra dalam pengenalan pola dan pengolahan gambar, serta dalam bidang bioinformatika untuk mengidentifikasi pola dalam data genetik dan penelitian biomedis.

Algoritma K-Means memiliki keunggulan dalam kesederhanaan dan efisiensi komputasi, namun juga memiliki beberapa keterbatasan, seperti sensitivitas terhadap inisialisasi pusat *cluster* dan keberadaan *outlier*. Berbagai penelitian telah dilakukan untuk mengatasi keterbatasan ini dan meningkatkan kinerja K-Means. Tinjauan literatur ini bertujuan untuk:

1. Menganalisis berbagai pendekatan dan implementasi algoritma K-Means dalam pengumpulan data *mining*.
2. Mengidentifikasi kekuatan dan kelemahan algoritma K-Means dalam konteks pengumpulan data *mining*.
3. Mengkaji modifikasi dan optimasi algoritma K-Means yang telah diusulkan dalam literatur.
4. Merumuskan arah penelitian lebih lanjut terkait penggunaan algoritma K-Means dalam pengumpulan data *mining*.

Melalui tinjauan literatur ini, diharapkan dapat diperoleh pemahaman yang lebih komprehensif mengenai peran serta potensi algoritma *K-Means* dalam proses *data mining*, khususnya dalam tahap pengelompokan data yang berperan penting dalam mengekstraksi pola dan wawasan yang tersembunyi. Selain itu, penelitian ini juga bertujuan untuk mengidentifikasi keunggulan serta keterbatasan algoritma *K-Means* dalam berbagai konteks aplikasi, sehingga dapat menjadi dasar bagi pengembangan metode yang lebih efektif dan efisien di masa depan. Dengan demikian, hasil dari tinjauan ini tidak hanya memberikan wawasan teoretis, tetapi juga membuka peluang inovasi dalam penerapan teknik *clustering* guna meningkatkan kualitas analisis data di berbagai bidang, seperti bisnis, kesehatan, dan teknologi informasi.

II. METODE PENELITIAN

Penelitian ini dilakukan dengan menelaah berbagai sumber jurnal ilmiah guna mengidentifikasi metode yang paling efektif dalam penerapan data mining. Proses ini melibatkan serangkaian tahapan yang sistematis untuk memastikan validitas dan relevansi informasi yang diperoleh. Tahapan tersebut mencakup penentuan kriteria kelayakan sebagai dasar seleksi metode, identifikasi sumber informasi yang kredibel, pemilihan literatur yang relevan dengan topik penelitian, pengumpulan data dari berbagai referensi yang telah diseleksi, serta pemilihan item data yang akan digunakan dalam analisis. Dengan pendekatan ini, penelitian dapat menghasilkan kesimpulan yang lebih akurat dan dapat diterapkan secara optimal dalam implementasi data mining.

Data dalam penelitian ini dikumpulkan secara manual dengan memanfaatkan aplikasi *Harzing Publish or Perish* untuk mencari jurnal yang relevan. Pencarian dilakukan dengan memasukkan kata kunci "Implementasi Data Mining," yang menghasilkan total 411.200 jurnal dari tiga sumber utama, yaitu Google Scholar, Semantic Scholar dan Crossref. Dari jumlah tersebut, dilakukan proses seleksi awal berdasarkan relevansi judul dan abstrak, sehingga tersaring 125 jurnal yang dianggap memenuhi kriteria sebagai kandidat referensi untuk menjawab pertanyaan penelitian. Selanjutnya, melalui tahap penyaringan lebih lanjut yang mempertimbangkan kesesuaian dengan topik dan kualitas penelitian, hanya 30 artikel yang dinyatakan memenuhi seluruh kriteria yang telah ditetapkan. Rincian data yang telah dikumpulkan disajikan dalam Tabel 1.

Tabel 1. Pengumpulan Data

Sumber	Studi Ditemukan Implementasi Data Mining	Kandidat	Terpilih
Crossef	1.000	28	6
Google Scholar	20.200	50	13
Sematic Scholar	390.000	47	11
Total	411.200	125	30

III. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengidentifikasi metode yang paling efektif dalam mengimplementasikan teknik *data mining*, dengan mempertimbangkan berbagai pendekatan yang telah dikaji dalam literatur serta praktik terbaik yang diterapkan di berbagai bidang. Melalui analisis komprehensif terhadap metode yang tersedia, penelitian ini mengevaluasi keunggulan dan keterbatasan masing-masing teknik guna menentukan pendekatan yang paling sesuai untuk diterapkan dalam konteks tertentu. Berdasarkan tujuan tersebut, hasil penelitian yang diperoleh telah dirangkum dan disajikan dalam Tabel 2 di bawah ini, yang memberikan gambaran jelas mengenai efektivitas setiap metode yang telah diuji dan dianalisis.

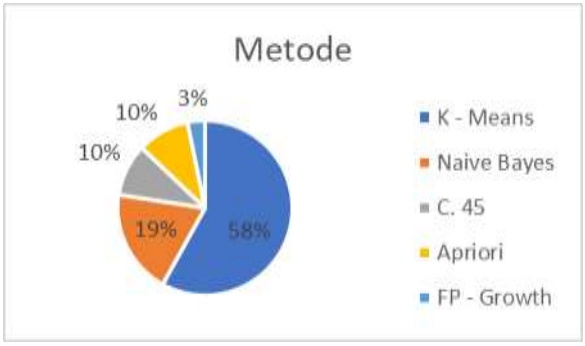
Tabel 2. Sumber Publikasi

No	Judul	Tahun	Jenis	Aspek
1	Implementasi Teknik	2024	Jurnal	Pariwisata
2	Analisis data mining ...	2021	Jurnal	Kesehatan
3	Analisis data mining ...	2022	Jurnal	Media Sosial
4	Analisis perbandingan ...	2021	Jurnal	Kesehatan
5	Analisis prediksi ...	2020	Jurnal	Pendidikan
6	Analisis data mining ...	2024	Jurnal	Pendidikan
7	Analisis data mining ...	2025	Jurnal	Bisnis
8	Penerapan data mining ...	2023	Jurnal	Sosial
9	Implementasi algoritma ...	2021	Jurnal	Bisnis
10	Implementasi data mining ...	2024	Jurnal	Pendidikan
11	Implementasi data mining ...	2024	Jurnal	Pendidikan
12	Implementasi data mining ...	2023	Jurnal	Sosial
13	Implementasi data mining ...	2023	Jurnal	Sosial
14	Implementasi data mining ...	2023	Jurnal	Sosial
15	Implementasi data mining ...	2024	Jurnal	Kesehatan
16	Implementasi data mining ...	2020	Jurnal	Kesehatan
17	Implementasi data mining ...	2024	Jurnal	Pendidikan
18	Implementasi data mining ...	2023	Jurnal	Pendidikan
19	Implementasi data mining ...	2024	Jurnal	Pendidikan
20	Implementasi data mining ...	2024	Jurnal	Pendidikan
21	Implementasi data mining ...	2022	Jurnal	Sosial
22	Implementasi data mining ...	2021	Jurnal	Sosial
23	Implementasi data mining ...	2021	Jurnal	Pendidikan
24	Implementasi data mining ...	2024	Jurnal	Pendidikan
25	Implementasi data mining ...	2024	Jurnal	Biotik
26	Implementasi data mining ...	2023	Jurnal	Pendidikan
27	Implementasi data mining ...	2020	Jurnal	Pendidikan
28	Implementasi data mining ...	2023	Jurnal	Sosial
29	Implementasi data mining ...	2021	Jurnal	Pendidikan
30	Implementasi data mining ...	2024	Jurnal	Teknologi

Tabel 3 menyajikan hasil penelitian yang telah dilakukan, menunjukkan metode yang paling umum digunakan dalam implementasi *data mining*. Dari analisis yang dilakukan, diketahui bahwa metode *klasifikasi* merupakan teknik yang paling banyak diterapkan dalam penelitian untuk mengidentifikasi pendekatan yang paling dominan digunakan dalam berbagai studi sebelumnya. Keunggulan metode *klasifikasi* terletak pada kemampuannya dalam meng-elompokkan data secara sistematis berdasarkan pola dan karakteristik tertentu, menjadikannya teknik yang potensial untuk terus berkembang dalam penelitian lanjutan. Jika dipelajari lebih lanjut, metode ini dapat dieksplorasi untuk menemukan variasi yang lebih efektif dalam meningkatkan akurasi dan efisiensi analisis data. Dengan pendekatan ini, penelitian di masa depan diharapkan dapat mengoptimalkan teknik *klasifikasi* dalam berbagai aplikasi, termasuk dalam deteksi berita hoaks secara profesional, sehingga dapat berkontribusi dalam mengurangi penyebaran informasi yang tidak valid di dunia digital.

Tabel 3. Daftar Penulis dan Metode

No	Penulis	Metode
1	S Wulandari, R Astuti, FM Basysyar [1]	K - Means
2	Z Nabila, AR Isnain, P Permata [2]	K – Means
3	F Prasetya, F Ferdiansyah [3]	Naïve bayes
4	F Solikhah, M Febianah, AL Kamil, WA Arifin, SJS Tyas [4]	Naïve bayes dan C4.5
5	L Setiyani, M Wahidin, D Awaludin, S Purwani [5]	Naïve bayes
6	N Ameliana, N Suarna, W Prihartono [6]	K – means
7	AOS Ayu, M Iqbal[7]	K – means
8	E Saputra, R Fauzi [8]	FP – growth
9	I. Kurniawan, R. H. Bhakti, B. Irawan [9]	K - means
10	N. Rahmawati, Zaehol Fatah [10]	Apriori
11	Willy Prihartono, Edi Tohidi, Ihsan Ahmad Fauzi, Raditya Danar Dana [11]	K – means
12	Arisman Waruwu, Milfa Yetri, F. Setiawan [12]	K – means
13	Amellia Veronica Agustin, A. Voutama [13]	Naïve bayes
14	Nurdin Nurdin, Cindy Cika Pradita, Fadlisyah Fadlisyah [14]	Apriori
15	Nurul Qolbi Rahmawati, Zaehol Fatah[15]	Apriori
16	Endang Etriyanti, Dedy Syamsuar, Yesi Novaria Kunang[16]	C4.5
17	Dimas Abisono Punkastyo, Fajar Septian, Ari Syaripudin[17]	Naïve bayes
18	Arisman Waruwu, Milfa Yetri, Feri Setiawan[18]	K – Means
19	Erna Nurliana, Bambang Irawan, Agus Bahtiar[19]	K – Means
20	Rini Astuti[20]	K – Means
21	Nisriina Nuur Hasanah, Agus Sidiq Purnomo[21]	K – Means
22	Haryani, Dicky Nofriansyah, Ita Mariami[22]	K – Means
23	Cici Armayani, Achmad Fauzi, Hermansyah Sembiring[23]	K – Means
24	Novitaria Manullang, Rahmat Widia Sembiring, Indra Gunawan, Iin Parlina, Irawan[24]	C4.5
25	Indra Agung Kurniawan, R M Herdian Bhakti, Bambang Irawan[25]	K – Means
26	Arisman Waruwu, Milfa Yetri, Feri Setiawan[26]	K – Means
27	Triase, Samsudin[27]	K – Means
28	Affani Putri Riyandoro, Apriade Voutama, Yuyun Umaidah[28]	K – Means
29	Amat Damuri, Umbar Riyanto, Hengki Rusdianto, Mohammad Aminudin[29]	Naïve bayes
30	Putri Pratiwi Pane, YusufRamadhanNasution, Mhd. Furqan[30]	K – Means



Gambar 1. Metode yang paling banyak digunakan dalam implementasi data mining

Berdasarkan data yang disajikan dalam Tabel 3 dan Gambar 1, diperoleh hasil bahwa algoritma *K-Means* merupakan metode yang paling efektif serta paling banyak digunakan dalam implementasi *data mining*. Hal ini menunjukkan bahwa teknik *clustering* dengan *K-Means* memiliki keunggulan dalam mengelompokkan data secara efisien berdasarkan karakteristik yang serupa, sehingga banyak diterapkan dalam berbagai bidang penelitian dan industri. Hasil survei ini dapat menjadi landasan bagi penelitian selanjutnya untuk mengeksplorasi pendekatan baru dalam implementasi *data mining*, baik melalui optimalisasi algoritma *K-Means* maupun pengembangan metode hibrida yang mengombinasikan teknik *clustering* dengan pendekatan lain guna meningkatkan akurasi dan efisiensi analisis data. Dengan demikian, penelitian di masa depan dapat berkontribusi dalam mengembangkan strategi yang lebih inovatif dan adaptif sesuai dengan kebutuhan analisis data yang semakin kompleks.

IV. SIMPULAN DAN SARAN

A. Simpulan

Penelitian ini secara khusus menyelidiki bagaimana cara terbaik menggunakan data mining dalam berbagai situasi. Dari berbagai teknik yang ada, clustering dengan algoritma K-Means menonjol sebagai metode yang sering digunakan dan efektif. Popularitas K-Means berasal dari kesederhanaannya dan kecepatannya dalam memproses data, membuatnya mudah diimplementasikan dalam berbagai aplikasi. Studi ini menemukan bahwa K-Means telah berhasil digunakan dalam berbagai bidang, seperti membagi pelanggan ke dalam kelompok-kelompok berdasarkan perilaku mereka, mengelompokkan minat konsumen, dan menganalisis tren penjualan.

Namun, penelitian ini juga menyoroti pentingnya menyadari keterbatasan K-Means. Salah satunya adalah sensitivitas terhadap pemilihan titik awal (centroid). Jika titik awal dipilih secara tidak tepat, hasil clustering bisa jadi kurang optimal. Selain itu, menentukan jumlah kelompok (cluster) yang paling sesuai juga bisa menjadi tantangan tersendiri.

Oleh karena itu, penelitian lebih lanjut diperlukan untuk mengatasi kelemahan-kelemahan ini. Misalnya, mengembangkan metode yang lebih kuat untuk memilih titik awal sehingga hasil clustering lebih konsisten. Selain itu, diperlukan cara yang lebih baik untuk menentukan jumlah kelompok yang optimal, sehingga clustering lebih akurat. Tantangan lainnya adalah bagaimana menangani situasi di mana sebuah data bisa masuk ke lebih dari satu kelompok (overlapping cluster). Teknik yang lebih canggih mungkin diperlukan untuk mengatasi masalah ini.

B. Saran

Pembahasan terkait penelitian ini masih sangat terbatas dan membutuhkan banyak masukan, saran untuk penulis selanjutnya adalah mengkaji lebih dalam dan secara komprehensif tentang Implementasi Data Mining Menggunakan Teknik Clustering dengan Metode K-Means.

DAFTAR RUJUKAN

Agustin, A. V., & Voutama, A. (2023). Implementasi data mining klasifikasi penyakit diabetes pada perempuan menggunakan Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika), 7*(2), 1002–1007. https://doi.org/10.36040/jati.v7i2.6808

Ameliana, N., Suarna, N., & Prihartono, W. (2024). Analisis data mining pengelompokkan UMKM menggunakan algoritma K-Means clustering di Provinsi Jawa Barat. *JATI (Jurnal Mahasiswa Teknik Informatika), 8*(3), 3261–3268. https://doi.org/10.36040/jati.v8i3.9655

Armayani, C., Fauzi, A., Sembiring, H., & Manajemen Informatika, P. (2021). Program sistem informasi STMIK Kaputama. *JIK, 5*(1).

Astuti, R., & Basysyar, F. M. (2024). Penerapan data mining clustering menggunakan metode K-Means pada data tindak kriminalitas di Polres Kabupaten Kuningan.

Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi data mining dengan algoritma Naïve Bayes untuk klasifikasi kelayakan penerima bantuan sembako. *JURIKOM (Jurnal Riset Komputer),* *8*(6), 219. https://doi.org/10.30865/jurikom.v8i6.3655

Dengan, U., & Clusstering, M. K. (2025). Analisis data mining terhadap data faktor perceraian di Sumatera. *[Nama Jurnal Tidak Dicantumkan], 4*, 214–221.

Fajrin, A. A., & Maulana, A. (2018). Penerapan data mining untuk analisis pol. *Kumpulan Jurnal Ilmu Komputer, 5*(1), 27–36.

Gunawan, I., Parlina, I., & Irawan, B. (2021). [Judul tidak lengkap]. *[Jurnal Tidak Disebutkan], 2*(2), 1–5. Retrieved from http://creativecommons.org/licences/by/4.0/

Hasanah, N. N., & Purnomo, A. S. (2022). Implementasi data mining untuk pengelompokan buku menggunakan algoritma K-Means clustering (Studi kasus: Perpustakaan Politeknik LPP Yogyakarta). *Jurnal Teknologi dan Sistem Informasi Bisnis,* *4*(2), 300–311. https://doi.org/10.47233/jteksis.v4i2.499

Haryani, D., Nofriansyah, D., & Mariami, I. (2021). Implementasi data mining untuk pengeelompokan buku di Perpustakaan Yayasan Nurul Islam Indonesia Baru dengan metode K-Means clustering. Retrieved from https://ojs.trigunadharma.ac.id/index.php/jct/index

Implementasi data mining menggunakan algoritme Naive Bayes classifier dan C4.5 untuk memprediksi kelulusan mahasiswa. (2020). *Telematika,* *13*(1), 56–67. https://doi.org/10.35671/telematika.v13i1.881

Irawan, B., Studi, P., Informatika, T., Teknik, F., Setiabudi, U. M., & Tengah, P. J. (2024). Implementasi data mining untuk mengukur prestasi siswa SD menggunakan metode K-Means clustering. *[Nama Jurnal Tidak Dicantumkan], 1*(2), 262–268.

Kurniawan, I. A., Bhakti, R. M. H., & Irawan, B. (2024). Implementasi data mining untuk mengukur prestasi siswa SD menggunakan metode K-Means clustering. *JCRD: Journal of Citizen Research and Development, 1*(2).

Kurniawan, Y. I. (2018). Perbandingan algoritma Naive Bayes dan C.45 dalam klasifikasi data mining. *Jurnal Teknologi Informasi dan Ilmu Komputer,* *5*(4), 455–464. https://doi.org/10.25126/jtiik.201854803

Napitupulu. (2017). UNIVERSITAS SUMATERA UTARA Poliklinik UNIVERSITAS SUMATERA UTARA. *Jurnal Pembangunan Wilayah & Kota, 1*(3), 82–91.

Nurdin, N., Pradita, C. C., & Fadlisyah, F. (2023). Implementasi data mining untuk menganalisis kategori kompetisi mahasiswa menggunakan algoritma apriori. *Sisfo: Jurnal Ilmiah Sistem Informasi,* *7*(1), 28. https://doi.org/10.29103/sisfo.v7i1.12104

Nurliana, E., Irawan, B., & Bahtiar, A. (2024). Implementasi data mining algoritma K-Means untuk klasifikasi penduduk miskin berdasarkan tingkat kemiskinan di Jawa Barat. *[Nama Jurnal Tidak Dicantumkan]*

Pane, P. P., Nasution, Y. R., & Furqan, M. (2024). Implementasi data mining dengan K-Means clustering untuk memprediksi pengadaan obat. *Journal of Computer System and Informatics (JoSYC),* *5*(2), 286–296. https://doi.org/10.47065/josyc.v5i2.4920

Prasetyo, R. B., Pranoto, Y. A., & Prasetya, R. P. (2023). Implementasi data mining menggunakan algoritma K-Means clustering penyakit pasien rawat jalan pada Klinik Dr. Atirah Desa Sioyong, Sulteng. *JATI (Jurnal Mahasiswa Teknik Informatika,* *7*(4), 2144–2151. https://doi.org/10.36040/jati.v7i4.7419

Pujianti, A., & Mulyawan, M. (2023). Implementasi data mining menggunakan metode K-Means clustering untuk menentukan status kematian bayi di Jawa Barat. *JATI (Jurnal Mahasiswa Teknik Informatika),* *7*(1), 459–463. https://doi.org/10.36040/jati.v7i1.6347

Punkastyo, D. A., Septian, F., & Syaripudin, A. (2024). Implementasi data mining menggunakan algoritma Naïve Bayes untuk prediksi kelulusan siswa. *[Nama Jurnal Tidak Dicantumkan]*

Rahmawati, N. Q., & Fatah, Z. (2024). Implementasi data mining terhadap data minat mahasiswa menggunakan metode apriori. *Jurnal Teknologi Kimia Unimal, 13*(2), 215–228. https://doi.org/10.29103/jtku.v13i2.19563

Rahmawati, N. Q., & Fatah, Z. (2024). Implementasi data mining terhadap data minat mahasiswa menggunakan metode apriori. *[Nama Jurnal Tidak Dicantumkan], 2*(November), 215–228.

Riyandoro, A. P., Voutama, A., Umaidah, Y., Waluyo, J. H. R., Timur Karawang, T., & Barat, J. (2023). Implementasi data mining clustering K-Means dalam menggolongkan beragam merek laptop. *[Nama Jurnal Tidak Dicantumkan]*

Setiyani, L., Wahidin, M., Awaludin, D., & Purwani, S. (2020). Analisis prediksi kelulusan mahasiswa tepat waktu menggunakan metode data mining Naïve Bayes: Systematic review. *Faktor Exacta, 13*(1), 35. https://doi.org/10.30998/faktorexacta.v13i1.5548

Triase, S. (2020). Implementasi data mining dalam mengklasifikasikan UKT (Uang Kuliah Tunggal) pada UIN Sumatera Utara Medan. *Jurnal Teknologi Informasi, 4*(2).

Waruwu, A., Yetri, M., & Setiawan, F. (2023). Implementasi data mining dalam mengelompokkan data penduduk kurang mampu menggunakan metode K-Means clustering. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD), 2*(6), 945. https://doi.org/10.53513/jursi.v2i6.8965

Waruwu, A., Yetri, M., & Setiawan, F. (2023). Implementasi data mining dalam mengelompokkan data penduduk kurang mampu menggunakan metode K-Means clustering. *Jurnal Sistem Informasi Triguna Dharma*. Retrieved from https://ojs.trigunadharma.ac.id/index.php/jsi

Zahra, N. A., Zidane, F. H., & Kuslaila, N. R. (2023). Analisis keamanan sistem informasi pada website PT Sentra Vidya Utama (Sevima) menggunakan metode OWASP. *Prosiding Seminar Nasional Teknologi dan Sistem Informasi, 3*(1), 384–393. https://doi.org/10.33005/sitasi.v3i1.564