

ALGORITMA K-MEANS CLUSTERING: SEBUAH STUDI LITERATUR

Pita sari ^{*1}, Efan², Riduan Syahri ³

¹Teknik Informatika, Institut Teknologi Pagar Alam, Sumatera Selatan
Email: ¹ pitasari248sr@gmail.com, ² efan@itpa.ac.id, ³syahririduan@gmail.com

Abstrak Tujuan utama penelitian ini adalah memanfaatkan teknologi *K-Means Clustering* dalam melakukan clustering data. Dengan melakukan tinjauan literatur dari tahun 2018 hingga 2023 melalui database jurnal populer seperti *Google Scholar*, penelitian ini bertujuan untuk mengidentifikasi pihak-pihak yang dapat memanfaatkan teknologi ini dan memperoleh solusi optimal dalam penerapannya. Metode SLR digunakan sebagai dasar untuk menemukan penelitian terkait sebelumnya dengan tujuan memperoleh pemahaman yang lebih mendalam tentang algoritma *K-Means Clustering*. Diharapkan bahwa hasil dari penelitian ini akan memberikan wawasan yang komprehensif bagi mereka yang tertarik dalam penerapan teknologi *K-Means* dalam analisis data.

Kata kunci: *K-Means Clustering; Sistematic Literature Review.*

K-MEANS CLUSTERING ALGORITHM: A LITERATUR STUDY

Abstract The main objective of this research is to utilize *K-Means Clustering* technology in clustering data. By conducting a literature review from 2018 to 2023 through popular journal databases such as *Google Scholar*, this research aims to identify parties who can utilize this technology and obtain optimal solutions in its application. The SLR method is used as a basis for finding previous related research with the aim of gaining a deeper understanding of the *K-Means Clustering* algorithm. It is expected that the results of this research will provide comprehensive insights for those interested in the application of *K-Means* technology in data analysis

Keywords: *K-Means Clustering; Systematic Literature Review.*

1. PENDAHULUAN

Proses data mining adalah tindakan mencari pola atau informasi menarik dalam data yang telah dipilih dengan menggunakan beragam teknik, metode, dan algoritma. [1]. *K-Means Clustering* adalah suatu metode analisis data yang menggunakan teknik data mining. Metode ini beroperasi secara unsupervised atau tanpa supervisi, dan termasuk dalam kategori pengelompokan data dengan pendekatan partisi [1]. Algoritma klasterisasi *K-Means* merupakan sebuah metode yang digunakan untuk mengelompokkan data berdasarkan titik pusat klaster terdekat. Algoritma ini bertujuan utama untuk memaksimalkan kesamaan data dalam satu klaster dan meminimalkan kesamaan data antar klaster. Untuk mencapai tujuan tersebut, fungsi jarak digunakan sebagai ukuran kemiripan dalam klaster. Dengan melakukan perhitungan jarak terpendek antara data dan titik centroid, kita dapat memperoleh pemaksimalan kemiripan data dalam klaster tersebut.

Pengelompokan data, yang sering disebut juga sebagai clustering adalah teknik yang kerap digunakan untuk mengelompokkan data menjadi kelompok-kelompok yang terpisah. Tujuan utamanya

adalah memisahkan dan mengelompokkan data yang memiliki kesamaan atau kemiripan dalam sifat atau karakteristik di antara data yang ada dalam dataset. Proses clustering digunakan untuk membagi data ke dalam kelas atau cluster berdasarkan tingkat kesamaannya [1]. Clustering atau klasifikasi adalah suatu teknik yang digunakan untuk membagi data menjadi beberapa kelompok berdasarkan kesamaan yang telah ditentukan sebelumnya [2].

Metode studi pustaka yang disebut *Systematic Literature Review* digunakan untuk melakukan evaluasi, klasifikasi, dan pengembangan landasan teori terhadap penelitian sebelumnya yang relevan dengan topik penelitian [3]. Metode penelitian yang dikenal sebagai *Systematic Literature Review* digunakan untuk melakukan evaluasi, klasifikasi, dan pengembangan dasar teori terhadap penelitian sebelumnya yang relevan dengan topik penelitian [4] SLR memegang peranan yang sangat penting dalam menentukan agenda penelitian, terutama dalam disertasi atau tesis, dan juga merupakan komponen yang krusial dalam aplikasi hibah penelitian. Tujuan utama dari *Systematic Literature Review* adalah untuk menggabungkan hasil-hasil penelitian primer

sehingga dapat menyajikan fakta yang lebih komprehensif dan seimbang.

Untuk lebih memahami algoritma K-means Clustering, Penelitian ini melakukan kajian pustaka dengan memanfaatkan Metode Tinjauan Literatur Sistematis (SLR). Metode SLR dapat memberikan kontribusi dalam menemukan solusi dengan melakukan review terhadap penelitian-penelitian terkait yang telah dilakukan sebelumnya. Tujuan utama dari penelitian ini adalah untuk menyajikan temuan yang komprehensif dan memberikan pandangan yang lebih luas mengenai aplikasi dan manfaat Algoritma K-Means Clustering dalam berbagai bidang. Untuk mencapai hasil yang optimal, penelitian dilakukan dengan memanfaatkan literatur yang telah diterbitkan dalam database jurnal terkemuka seperti Google Scholar. Rentang waktu yang digunakan dalam penelitian ini adalah dari tahun 2018 hingga 2023.

2. METODE PENELITIAN

SLR (Systematic Literature Review) merupakan sebuah penelitian sekunder dalam rangka memetakan, mengidentifikasi, mengevaluasi secara kritis, mengkonsolidasikan, dan mengumpulkan hasil-hasil dari studi utama yang telah dilakukan pada topik penelitian yang spesifik, dilakukanlah suatu upaya yang bertujuan untuk mencapai hal tersebut [5]. Salah satu cara yang umum digunakan Untuk memperoleh jawaban, salah satu metode yang dapat digunakan adalah dengan melakukan tinjauan pustaka berdasarkan studi terkait sebelumnya. Sistematis Literature Review (SLR) bertujuan untuk mengumpulkan dan merangkum penelitian-penelitian sebelumnya, mengidentifikasi kesenjangan yang perlu diisi antara penelitian sebelumnya dan saat ini, serta menyusun laporan atau sintesis yang koheren dan membuat kerangka kerja penelitian, diperlukan upaya yang sistematis dan terstruktur.

Tujuan dari studi literatur ini adalah untuk menyajikan temuan yang komprehensif dan memberikan pandangan yang lebih luas mengenai aplikasi dan manfaat Algoritma K-Means Clustering dalam berbagai bidang. Agar dapat mencapai hasil yang optimal, penelitian dilakukan dengan menggunakan literatur yang telah diterbitkan dalam database jurnal populer seperti Google Scholar dari tahun 2018 hingga 2023.

2.1 Research Question

Untuk menjaga fokus tinjauan literatur, tujuan dari pertanyaan penelitian adalah memastikan agar tinjauan literatur tetap terfokus. Dengan demikian, pencarian data yang diperlukan dapat dilakukan dengan lebih lancar. Pertanyaan penelitian untuk penelitian ini dapat ditemukan pada Tabel 1.

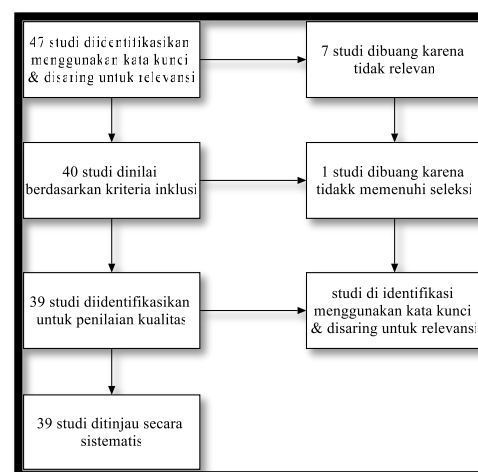
Table 1 Research Question

3.2. Article Quality Assesment

Penelitian ini dilaksanakan dengan tujuan untuk melakukan pencarian literatur atau studi literatur

ID	Research Question	Motivasi
RQ01	Bagaimana konsep Algoritma K-Means Clustering diterapkan dalam berbagai konteks?	Identifikasi topik riset mengenai Algoritma K-means Clustering dan penerapannya dalam berbagai konteks

secara sistematis, peneliti menggunakan metode pengumpulan data yang meliputi identifikasi, evaluasi, interpretasi, literatur, dan laporan yang tersedia. Peneliti melakukan tinjauan, mengidentifikasi jurnal yang relevan, dan mengikuti langkah-langkah yang ditunjukkan oleh prosedur dalam setiap proses. secara sistematis. Setelah survei selesai, Fokus utama dari penelitian ini adalah mengenai hasil belajar, metode membaca, dan pendekatan terbuka. Sebanyak 40 artikel jurnal yang diterbitkan antara 2018-2023 dikumpulkan oleh peneliti sebagai bahan penelitian. Peneliti juga mengumpulkan artikel majalah dengan kata kunci "Algoritma k-means Clustering" untuk penelitian ini. Dalam bagian ini, diharapkan agar peneliti dapat menjelaskan metode yang digunakan dalam penelitiannya dengan memanfaatkan gambar, diagram, dan diagram alur.



Gambar 1 Tahapan Studi Literatur

3. HASIL DAN PEMBAHASAN

3.1 Masalah

Masalah yang ditemukan mengenai Algoritma K-means Clustering di terapkan ke dalam 4 aspek yakni aspek Sosial, Pemasaran produk, Pertanian dan Algoritma.

1. Masalah yang berkaitan aspek sosial

Terkait aspek sosial, Algoritma K-means di terapkan pada permasalahan tingkat kemiskinan, ketidakmerataan pendapatan, dan prioritas penerima bantuan, seperti yang ditunjukkan pada table 2

Table 1 Masalah terkait aspek sosial

Masalah	Sumber
Tingkat kemiskinan	[11] [3] [4] [5] [6] [7] [9] [36]
Ketidakmerataan pendapatan	[2]
Prioritas penerima bantuan	[11] [18] [32].

Tingkat kemiskinan, dengan tingkat kemiskinan yang cukup tinggi diperlukan Upaya untuk mengidentifikasi dan klasifikasi data kemiskinan [11]. Kurangnya kesempatan kerja yang memadai juga berkontribusi terhadap tingkat kemiskinan [3]. Angka kemiskinan di Indonesia terus meningkat dikarenakan masyarakat tidak memiliki kemampuan untuk memenuhi kebutuhan dasar mereka, seperti pakaian, makanan, dan tempat tinggal [7]. Pengelompokan data keluarga miskin dengan tepat dan efisien sangat diperlukan untuk mengetahui kondisi kemiskinan [5]. Dalam rangka mengatasi masalah kemiskinan, penting bagi pemerintah untuk mengetahui tingkat kemiskinan yang ada di setiap daerah. Dengan memahami wilayah-wilayah yang memiliki tingkat kemiskinan yang tinggi, sedang, atau rendah, pemerintah dapat menentukan prioritas dalam upaya mengatasi masalah kemiskinan [9]. Pentingnya peran data mining dalam pengolahan data berskala besar di berbagai bidang kehidupan, seperti industri, keuangan, cuaca, ilmu, dan teknologi tidak dapat diabaikan. Terdapat beberapa metode yang dapat digunakan dalam data mining, seperti klasifikasi, regresi, seleksi variabel, dan clustering [36]. Metode analisis pengelompokan adalah metode analisis yang bertujuan untuk memetakan atau mengkategorikan objek data berdasarkan kesamaan karakteristik di antara objek data tersebut [6]. Dengan menggunakan K-Means, yang merupakan salah satu metode pengelompokan yang sering dipakai [4].

Ketidakmerataan pendapatan, Untuk mengatasi ketidakmerataan pendapatan di setiap daerah, diperlukan pendekatan yang berbeda-beda antara wilayah sesuai dengan kondisi yang ada di setiap kabupaten/kota. Salah satu pendekatan yang dapat diterapkan adalah dengan mengelompokkan ciri-ciri wilayah berdasarkan indikator kemiskinan menggunakan metode pengelompokan K-means Clustering. Dengan menggunakan pendekatan ini,

kita dapat mengidentifikasi pola dan kelompok-kelompok yang ada dalam data kemiskinan, sehingga dapat dilakukan penanganan yang lebih spesifik dan efektif di setiap wilayah [2].

Prioritas penerima bantuan, banyak bantuan yang diberikan kepada penduduk miskin di Indonesia memiliki prioritas yang berbeda. Contohnya adalah Jaminan Kesehatan Nasional (JKN), Kartu Indonesia Sehat (KIS), Kartu Indonesia Pintar (KIP), dan Kartu Keluarga Harapan (KKH) [11]. Data yang sering diabaikan dalam validasi dapat menyebabkan ketidakakuratan data [18]. Metode analisis cluster merupakan suatu pendekatan multivariat yang digunakan untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimiliki oleh masing-masing objek tersebut [32].

2. Masalah yang berkaitan pemasaran produk

Terkait Pemasaran produk, Algoritma K-means di terapkan pada permasalahan penjualan produk, dan persediaan stok.

Table 3 Masalah terkait pemasaran produk

Masalah	Sumber
Penjualan produk	[13] [16] [27] [29] [39] [40] [35]
Persediaan stok	[10] [24] [30]

Penjualan produk, kebutuhan untuk menganalisis penjualan produk dengan tujuan mengembangkan strategi pemasaran yang lebih efektif, menentukan produk-produk yang paling diminati dan yang paling tidak diminati perlu dilakukan [16]. Dengan memanfaatkan data transaksi, kita dapat melakukan pengelompokan untuk menganalisis produk yang memiliki tingkat penjualan yang tinggi dan rendah di pasaran. [35]. Keakuratan estimasi penjualan memiliki pengaruh yang signifikan terhadap penjualan [40]. Data transaksi penjualan merupakan fakta yang dapat diolah dan dianalisis untuk menghasilkan informasi yang berharga bagi kegiatan bisnis. [29]. Salah satu metode analisis Data Mining yang umum digunakan adalah analisis kelompok atau yang lebih dikenal dengan Clustering. [39]. Menentukan jumlah cluster yang optimal untuk sebuah set data adalah masalah penting dalam algoritma clustering tertentu, terutama k-means yang terkenal dan algoritma serupa [27]. K-Means merupakan salah satu teknik pengelompokan data non-hirarki yang membagi data ke dalam satu atau lebih cluster atau kelompok [13].

Persediaan stok, Persediaan mengacu pada sekumpulan material yang terdiri dari bahan baku, barang dalam proses, dan barang jadi. Sementara itu, pengendalian persediaan adalah tindakan untuk menjaga jumlah persediaan agar tetap berada pada tingkat yang diinginkan. [24]. Data mining merupakan serangkaian proses yang digunakan untuk menemukan pola, keterkaitan, dan nilai tambah dari data dan informasi yang memiliki volume yang besar. Data mining memiliki tujuan untuk menemukan pola dan mengolah data sehingga menghasilkan informasi yang dapat dimengerti dan berguna. Metode data mining ini didukung oleh ilmu statistik dan matematika untuk menganalisis data secara mendalam untuk mencapai tujuan tersebut [30]. Mengidentifikasi pola dan tren dalam data penjualan serta mengelompokkan barang-barang berdasarkan tingkat persediaan stoknya perlu dilakukan agar dapat mengambil keputusan yang lebih baik dalam mengatur persediaan stok untuk itu perlu mengimplementasikan metode K-means Clustering [10].

3. Masalah yang terkait aspek pertanian

Terkait aspek pertanian, Algoritma K-means di terapkan pada produksi pangan, dan lahan kritis.

Table 2 masalah terkait pertanian

Masalah	Sumber
Produksi pangan	[28]
Lahan kritis	[14,33]

Produksi pangan, Sangat menguntungkan apabila dilakukan analisis klaster terhadap seluruh provinsi di Indonesia berdasarkan jenis produksi bahan pangan. Dalam setiap provinsi, bahan pangan utama dapat dihitung berdasarkan persentase hasil produksi dari padi, kacang tanah, kedelai, jagung, dan ubi kayu. Metode K-Means digunakan untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan produksi bahan pangan [28].

Lahan kritis, Penurunan fungsi lahan yang kritis dapat berdampak negatif terhadap tata air dan sumber daya tanah. Dampak ini dapat berpengaruh terhadap kehidupan sosial dan ekonomi masyarakat yang bergantung pada lahan tersebut. Oleh karena itu, diperlukan pengelompokan provinsi untuk mengetahui Tingkat penyebaran lahan yang kritis masih tinggi. Salah satu cara untuk mengatasi masalah ini adalah dengan menerapkan metode Data Mining menggunakan algoritma K-Means. Metode

ini dapat membantu dalam mengidentifikasi pola dan kelompok lahan kritis yang memiliki karakteristik serupa. Dengan demikian, langkah-langkah yang sesuai dapat diambil untuk mengelola dan memulihkan lahan yang mengalami kerusakan [14]. Metode K-Means Clustering adalah sebuah algoritma dalam Data Mining yang digunakan untuk mengelompokkan data ke dalam beberapa cluster berdasarkan kesamaan karakteristik. Data yang memiliki karakteristik yang serupa akan dikelompokkan ke dalam satu cluster, sedangkan data yang memiliki karakteristik yang berbeda akan dikelompokkan ke dalam cluster yang berbeda pula [33].

4. Masalah terkait Algoritma K-Means Clustering

Table 5 Masalah terkait Algoritma

Masalah	Sumber
Optimasi algoritma	[19] [26]
Penerapan algoritma	[8] [12] [15] [17] [20] [21] [22] [23] [31] [34] [37] [38]

Optimasi algoritma, Dalam optimasi algoritma, metode elbow digunakan untuk mengoptimalkan algoritma dengan tujuan menentukan nilai SSE (Sum of Square Error). Metode Elbow digunakan untuk menghasilkan informasi dalam menentukan jumlah cluster terbaik dengan memperhatikan persentase hasil perbandingan antara jumlah cluster (k) yang akan membentuk siku pada suatu titik [19]. Untuk menentukan jumlah cluster yang optimal, dapat dilakukan dengan memperhatikan persentase hasil perbandingan antara jumlah cluster (k) yang akan membentuk siku pada suatu titik [26].

Penerapan Algoritma, Metode pengelompokan dapat diterapkan dalam bidang pemasaran guna mengatasi berbagai permasalahan yang muncul [17]. Dalam rangka melakukan klasterisasi atau pengelompokan, diperlukan metode yang dapat secara otomatis mengolah data peraturan daerah [20]. Salah satu manfaat dari algoritma K-Means adalah kemampuannya dalam memberikan tingkat akurasi yang tinggi terhadap ukuran objek. Keunggulan ini menjadikan algoritma ini lebih terukur dan efisien dalam mengolah objek dalam jumlah yang besar [23]. Analisis klaster adalah salah satu teknik dalam analisis statistika multivariat [31]. Terkadang, para peneliti dan analis berusaha untuk menemukan cara yang simpel dalam menggambarkan pola dan tren yang terdapat dalam data [37]. Pengelompokan memiliki peran yang lebih dari sekadar mengelompokkan dokumen atau benda [38]. Tujuan dari pengelompokan data ini adalah untuk mengurangi fungsi objektif yang telah ditetapkan

dalam proses pengelompokan [34]. Algoritma K-Means memanfaatkan nilai rata-rata guna menentukan pusat data (centroid) [15]. Ada dua jenis pengelompokan yang sering digunakan dalam proses pengelompokan data, yaitu pengelompokan hierarki dan pengelompokan non-hierarki. Salah satu metode pengelompokan non-hierarki yang sering dipakai adalah K-Means, yang bertujuan untuk membagi data yang ada menjadi satu atau lebih kelompok [21]. K-Means merupakan salah satu teknik pengelompokan yang umum digunakan dalam penerapan data mining untuk mengklaster atau mengelompokkan data menjadi beberapa kelompok data [12]. Metode K-Means merupakan suatu teknik yang digunakan untuk mengelompokkan data ke dalam beberapa bagian. Metode ini berfokus pada jarak antara data dan membaginya menjadi beberapa kelompok. Namun, perlu diperhatikan bahwa metode K-Means hanya dapat digunakan pada atribut numerik. K-Means termasuk dalam jenis clustering partisi yang memisahkan data ke dalam k daerah yang berbeda [8]. Metode pengelompokan bertujuan untuk mengelompokkan sejumlah data atau objek ke dalam cluster atau grup dengan tujuan agar setiap cluster berisi data yang memiliki kesamaan sebanyak mungkin [22].

4. SIMPULAN

Kesimpulan dari penelitian literatur ini adalah bahwa algoritma K-Means memiliki berbagai aplikasi yang relevan dan dapat memberikan kontribusi penting dalam menyelesaikan permasalahan di berbagai bidang. Dalam setiap konteks, penggunaan algoritma ini memberikan pemahaman yang lebih mendalam terhadap data dan mendukung pengambilan keputusan yang lebih cerdas.

5. DAFTAR PUSTAKA

- [1] F. Yunita, "Yunita, Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru (Studi Kasus : Universitas Islam Indragiri) 1 238," vol. 7, no. September, pp. 238–249, 2018.
- [2] Y. Darmi and A. Setiawan, "PENERAPAN METODE CLUSTERING K-MEANS DALAM," vol. 12, no. 2, pp. 148–157, 2016.
- [3] A. A. Salim, M. Akbar, M. A. Modestio, and ..., "Systematic Literature Review: Analisis Penerapan Algoritma YOLO Dalam Mendeteksi Objek Jenis Makanan Ringan," *JRIIN J. Ris. ...*, vol. 1, no. 2, pp. 498–503, 2023, [Online]. Available: <http://jurnalmahasiswa.com/index.php/jriin/article/view/200%0Ahttps://jurnalmahasiswa.com/index.php/jriin/article/download/200/248>
- [4] C. Kamila, "Systematic Literature Review: Penggunaan Algoritma K-Means Untuk Clustering di Indonesia dalam Bidang Pendidikan," *Intech*, vol. 2, no. 1, pp. 19–24, 2021, doi: 10.54895/intech.v2i1.866.
- [5] F. Rozi, "Systematic Literature Review pada Analisis Prediktif dengan IoT : Tren Riset , Metode , dan Arsitektur," *J. Sist. Cerdas 2020 Vol*, vol. 03, no. 01, pp. 43–53, 2020.
- [6] M. R. Nahjan, N. Heryana, A. Voutama, F. I. Komputer, U. S. Karawang, and R. Miner, "IMPLEMENTASI RAPIDMINER DENGAN METODE CLUSTERING K-MEANS UNTUK ANALISA PENJUALAN PADA TOKO OJ CELL," vol. 7, no. 1, pp. 101–104, 2023.
- [7] Y. Hartati, S. Defit, and G. W. Nurcahyo, "Jurnal Informatika Ekonomi Bisnis Klasterisasi Bibit Terbaik Menggunakan Algoritma K-Means dalam Meningkatkan Penjualan," vol. 3, pp. 4–10, 2021, doi: 10.37034/infec.v3i1.56.
- [8] R. Hablum, "Clustering Hasil Tangkap Ikan Di Pelabuhan Perikanan Nusantara (Ppn) Ternate Menggunakan Algoritma K-Means," *JIKO (Jurnal Inform. dan Komputer)*, vol. 2, no. 1, pp. 26–33, 2019, doi: 10.33387/jiko.v2i1.1053.
- [9] S. Hajar, A. A. Novany, A. P. Windarto, A. Wanto, and E. Irawan, "Penerapan K-Means Clustering pada ekspor minyak kelapa sawit menurut negara tujuan," *Semin. Nas. Teknol. Komput. Sains 2020*, pp. 314–318, 2020.
- [10] K. Aprilia and F. Sembiring, "ANALISIS GARIS KEMISKINAN MAKANAN MENGGUNAKAN METODE ALGORITMA K-MEANS CLUSTERING," 2021.
- [11] D. N. Alfiansyah, V. Rahmayanti, S. Nastiti, and N. Hayatin, "Penerapan Metode K-Means pada Data Penduduk Miskin Per Kecamatan Kabupaten Blitar," vol. 4, no. 1, pp. 49–58, 2022.
- [12] R. Yuliani, "Infotek : Jurnal Informatika dan Teknologi Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur Infotek : Jurnal Informatika dan Teknologi Pendahuluan masalah kemiskinan belum bis," vol. 4, no. 1, pp. 39–50, 2021.
- [13] N. I. Febianto and N. D. Palasara, "Analisis Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018," vol. 08, no. September, pp. 130–140, 2019.
- [14] Y. Dwi and B. Winarno, "Implementasi Metode Mcquitty Linkage dan K-Means Clustering dengan Pendekatan Rule-of-Thumb pada Data Kemiskinan di Jawa Tengah," vol. 6, pp. 751–756, 2023.
- [15] R. Fikri, A. Mushardiyanto, M. Naufal, L. Banin, K. Maureen, and H. Patria,

- “Pengelompokan Kabupaten / Kota di Indonesia Berdasarkan Informasi Kemiskinan Tahun 2020 Menggunakan Metode K-Means Clustering Analysis,” vol. 1, no. November, pp. 190–199, 2021.
- [16] A. Bahauddin, A. Fatmawati, and F. Permata Sari, “Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means,” *J. Manaj. Inform. dan Sist. Inf.*, vol. 4, no. 1, pp. 1–8, 2021, doi: 10.36595/misi.v4i1.216.
- [17] D. Ananda, S. Rohimah, B. Susilo, D. Wulan, and A. Mustofa, “Implementation of K-Means in Grouping Birth Certificate Data in Indonesia Implementasi K-Means Dalam Pengelompokan Data Akta Kelahiran di Indonesia,” pp. 66–71, 2022.
- [18] Y. R. Sari *et al.*, “PENERAPAN ALGORITMA K- MEANS UNTUK CLUSTERING DATA KEMISKINAN PROVINSI BANTEN MENGGUNAKAN RAPIDMINER,” vol. 5, no. 2, pp. 192–198, 2020.
- [19] Z. Aras and Sarjono, “Analisis Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan Bedah Rumah Menggunakan Metode Clustering K-Means(Studi Kasus: Kantor Kecamatan Bahar Utara),” *J. Manaj. Sist. Inf.*, vol. 1, no. 2, pp. 159–170, 2016.
- [20] J. Informatika *et al.*, “Penerapan Algoritma Clustering K-Means Untuk Menentukan Prioritas Penerima Bantuan Dana Sosial PKH Di Kelurahan Kampung Singkep Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM),” vol. 3, no. April, pp. 342–350, 2023.
- [21] M. G. Sadewo, A. Eriza, A. P. Windarto, and D. Hartama, “Algoritma K-Means Dalam Mengelompokkan Desa / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Berdasarkan Provinsi,” pp. 754–761, 2019.
- [22] F. Putri, A. Hasibuan, S. Sumarno, and I. Parlina, “Penerapan K-Means pada Pengelompokan Penjualan Produk Smartphone,” vol. 1, no. 1, pp. 15–20, 2021, doi: 10.54259/satesi.v1i1.3.
- [23] M. Method, F. Indriyani, and E. Irfiani, “Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means (Clustering Sales Data at Outdoor Equipment Stores Using K-,” vol. 7, no. November, pp. 109–113, 2019.
- [24] H. Prastiwi, Jeny Pricilia, and Errisya Rasywir, “Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering,” *J. Inform. Dan Rekayasa Komputer(JAKAKOM)*, vol. 2, no. 1, pp. 141–148, 2022, doi: 10.33998/jakakom.2022.2.1.34.
- [25] A. Sani, “Penerapan metode k-means clustering pada perusahaan,” no. 353, pp. 1–7.
- [26] I. R. Mahartika and A. Wibowo, “Data Mining Klasterisasi dengan Algoritme K-Means untuk Pengelompokkan Provinsi Berdasarkan Konsumsi Bahan Bakar Minyak Nasional”.
- [27] C. Nicholas, “Mohon berikan umpan balik Mohon dukung repositori ScholarWorks @ UMBC dengan mengirim email ke scholarworks-group@umbc.edu dan beri tahu kami apa arti memiliki akses ke karya ini bagi Anda dan mengapa ini penting bagi Anda . Terima,” 2020.
- [28] T. Tendean, W. Purba, and M. Kom, “Analisis Cluster Provinsi Indonesia Berdasarkan Produksi Bahan Pangan Menggunakan Algoritma K-Means,” vol. 1, no. 2, pp. 5–11, 2020.
- [29] P. P. Siregar and Z. A. Siregar, “Penerapan Metode K-Means Dalam Mengelompokkan Persebaran Lahan Kritis Di Indonesia Berdasarkan Provinsi,” vol. 2, no. 4, pp. 145–151, 2022.
- [30] J. Informasi, “Prediksi Tingkat Kerugian Peternak Akibat Penyakit pada Sapi Menggunakan Algoritma K-Means Clustering,” vol. 3, 2021, doi: 10.37034/jidt.v3i1.87.
- [31] A. Febriyanto, S. Achmadi, and A. P. Sasmito, “Penerapan Metode K-Means Untuk Clustering Pengunjung Perpustakaan Itn Malang,” *J. Mhs. Tek. Inform.*, vol. 5, no. 1, pp. 61–70, 2021.
- [32] N. N. Hasanah and A. S. Purnomo, “Implementasi Data Mining Untuk Pengelompokan Buku Menggunakan Algoritma K-Means Clustering (Studi Kasus: Perpustakaan Politeknik LPP Yogyakarta),” *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 4, no. 2, pp. 300–311, 2022, doi: 10.47233/jteksis.v4i2.499.
- [33] R. P. K. Covid-, B. Kota, and K. Di, “PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS DALAM CLUSTERING,” vol. 9, no. 2, pp. 126–131, 2022, doi: 10.30656/jsii.v9i2.5127.
- [34] D. E. Putra, L. Muhelni, F. Teknik, U. Nahdlatul, and U. Sumatera, “Penerapan data mining pemasaran produk menggunakan metode clustering,” vol. 5, pp. 160–167, 2022, doi: 10.37600/tekinkom.v5i1.458.
- [35] N. A. Rakhmawati, D. B. Aletha, L. Putu, G. Widiastuti, A. Sidum, and L. Kaki, “Klasterisasi Peraturan Daerah di Seluruh Wilayah Jawa dengan Menggunakan Algoritma K-means,” vol. 09, pp. 242–246,

- 2020.
- [36] M. Terkait, “K-Means – Penerapan, Permasalahan dan Metode Terkait,” vol. 3, no. Pebruari, pp. 47–60, 2007.
 - [37] M. N. Mara and N. Satyahadewi, “PENGKLASIFIKASIAN KARAKTERISTIK DENGAN METODE K-MEANS CLUSTER ANALYSIS,” vol. 02, no. 2, pp. 133–136, 2013.
 - [38] A. Septianingsih, “ANALISIS K-MEANS CLUSTERING PADA PEMETAAN PROVINSI,” vol. 3, no. 1, pp. 224–241, 2022.
 - [39] M. Hasanah, S. Defit, G. W. Nurcahyo, T. Informatika, and F. I. Komputer, “Jurnal Sistim Informasi dan Teknologi Implementasi Algoritma K-Means untuk Klasterisasi Peserta Olimpiade Sains Nasional Tingkat SMA,” vol. 3, pp. 30–35, 2021, doi: 10.35134/jsisfotek.v1i3.6.
 - [40] A. N. Khomarudin, “Teknik Data Mining : Algoritma K-Means Clustering,” pp. 1–12, 2016.