

PENGELOMPOKAN DATA VARIAN PEKERJAAN DAN STATUS PERNIKAHAN PT DIKA MENGGUNAKAN ALGORITMA *K-MEANS CLUSTERING*

Aynun Nabila¹, Yudhistira Arie Wijaya²

¹ Komputerasi Akuntansi, STMIK IKMI Cirebon

² Sistem Informasi, STMIK IKMI Cirebon

Jalan Perjuangan No 10B, Karyamulya, Kec. Kesambi, Kota Cirebon

nabilaaynun870@gmail.com

ABSTRAK

PT Dika saat ini mengalami kesulitan yang signifikan dalam mengelola data transaksi penjualan di setiap cabangnya. Kesulitan ini muncul karena kompleksitas volume data yang terus meningkat seiring dengan perkembangan operasional bisnis. Setiap cabang turut berkontribusi pada pencatatan transaksi penjualan yang melibatkan berbagai elemen, mulai dari identifikasi produk hingga informasi pelanggan. Peningkatan jumlah cabang dan aktivitas transaksi secara keseluruhan membuat tugas pengelolaan data semakin rumit. Puncaknya, keberagaman format dan sumber data dari setiap cabang menambah kompleksitas tugas tersebut. Sebagai hasilnya, perusahaan mengalami kesulitan dalam mengintegrasikan, memproses, dan menganalisis data transaksi dengan efisien dan efektif. Tujuan dari penelitian ini adalah untuk menemukan solusi yang memungkinkan PT Dika untuk menganalisis data, meningkatkan strategi pemasaran, dan meningkatkan retensi pelanggan. Algoritma *K-Means*, dikenal efektif dalam pengelompokan data, diimplementasikan untuk memahami pola perilaku pelanggan dan meningkatkan manajemen data transaksional. DBI digunakan untuk menilai kualitas klastering dengan fokus pada homogenitas dan separabilitas cluster. Melalui penelitian ini, diharapkan dapat ditemukan solusi yang memungkinkan PT Dika untuk mengoptimalkan analisis data, meningkatkan strategi pemasaran, dan meningkatkan retensi pelanggan. Hasil dari implementasi algoritma *K-Means* pada data tersebut menunjukkan bahwa data dapat dikelompokkan menjadi tiga kelompok yang homogen berdasarkan karakteristik pekerjaan dan status pernikahan. Kelompok pertama terdiri dari karyawan yang belum menikah dan bekerja sebagai staf administrasi, kelompok kedua terdiri dari karyawan yang sudah menikah dan bekerja sebagai supervisor, dan kelompok ketiga terdiri dari karyawan yang sudah menikah dan bekerja sebagai manajer.

Kata kunci: PT Dika, data transaksional, algoritma *K-Means*, varian pekerjaan, status pernikahan, *Davies-Bouldin Index (DBI)*.

1. PENDAHULUAN

PT Dika saat ini mengalami kesulitan yang signifikan dalam mengelola data transaksi penjualan di setiap cabangnya. Kesulitan ini muncul karena kompleksitas volume data yang terus meningkat seiring dengan perkembangan operasional bisnis. Setiap cabang turut berkontribusi pada pencatatan transaksi penjualan yang melibatkan berbagai elemen, mulai dari identifikasi produk hingga informasi pelanggan [1]. Peningkatan jumlah cabang dan aktivitas transaksi secara keseluruhan membuat tugas pengelolaan data semakin rumit. Puncaknya, keberagaman format dan sumber data dari setiap cabang menambah kompleksitas tugas tersebut. Sebagai hasilnya, perusahaan mengalami kesulitan dalam mengintegrasikan, memproses, dan menganalisis data transaksi dengan efisien dan efektif.

Pada Penelitian ini, metode yang digunakan untuk mengelompokkan sebuah varian status pekerjaan dan status pernikahan dengan menggunakan algoritma *k-means* sebagai teknik pengelompokan. Algoritma *K-Means* adalah teknik penambangan data yang menggunakan metode pengelompokan atau pengelompokan untuk proses pemodelannya, tanpa pengawasan atau pembelajaran. Data diatur ke dalam partisi yang berbeda, dengan masing-masing kelompok menunjukkan kesamaan di antara

anggotanya sementara juga memiliki fitur berbeda yang membedakannya dari kelompok lain. Tujuannya adalah untuk meminimalkan perbedaan dalam cluster dan memaksimalkan disparitas sehubungan dengan cluster lain [2].

Penelitian yang dilakukan oleh Mohammad Syaefulloh pada tahun 2022 dengan jurnal yang berjudul “Clustering Kelompok Belajar Siswa Berdasarkan Hasil Ujian Sekolah Menggunakan Algoritma *K-Means*”. Algoritma *K-Means*, sebagai metode klastering yang telah terbukti efektif, mampu menghasilkan cluster dengan atribut yang tidak hanya menguntungkan tetapi juga terpuji. Pendekatan ini menjadi semakin relevan dalam konteks analisis data, di mana keberhasilan pengelompokan dapat memainkan peran kritis dalam pemahaman pola-pola yang mendasari data. Dengan mengidentifikasi karakteristik yang serupa, Algoritma *K-Means* membentuk *cluster-cluster* yang memberikan wawasan mendalam terhadap sifat dan tren yang mungkin tidak terlihat secara langsung [3]. Pada penelitian selanjutnya yang dilakukan oleh Masjunaedi dan Nana Suarna pada tahun 2023 dengan judul jurnal “Analisa Penerapan Metode Clustering *K-Means* Untuk Pengelompokan Data Transaksi Konsumen (Studi Kasus: Cv. Mitra Indexindo Pratama)” yaitu pemanfaatan pengelompokan *K-*

Means dalam konteks analisis data transaksi pelanggan merupakan pendekatan yang sangat bermanfaat dalam mengeksplorasi pola dan karakteristik yang mendasari perilaku pelanggan. Algoritma *K-Means*, sebagai metode clustering yang terkenal, memberikan kemampuan untuk mengelompokkan data transaksi pelanggan ke dalam cluster berdasarkan kemiripan atribut tertentu. Dengan menggunakan pendekatan ini, penelitian atau aplikasi bisnis dapat mendapatkan pemahaman yang lebih mendalam tentang segmen pelanggan yang mungkin memiliki preferensi atau kecenderungan serupa dalam melakukan transaksi [4].

Pada penelitian yang dilakukan Permadi dan Wijaya pada tahun 2023 dengan judul “Pengelompokan Dataset Bus Menggunakan Algoritma K-Means “yaitu pemanfaatan *Davies Bouldin Index* (DBI) sebagai parameter evaluasi menjadi suatu pendekatan yang sangat penting dan strategis dalam menilai kualitas pengelompokan data. DBI merupakan metrik evaluasi yang mengukur seberapa baik dan sejauh mana pembentukan *cluster-cluster* oleh suatu algoritma klustering, seperti *K-Means*, dapat memberikan hasil yang optimal. Dengan menggunakan DBI, penelitian atau implementasi algoritma klustering dapat mengukur tingkat keseragaman antara *cluster*, sekaligus mempertimbangkan tingkat kemiripan antar-*cluster* [5]. Penelitian yang dilakukan oleh Moh. Nurdayat dan Nana Suarna pada tahun 2023 yang berjudul “Analisa Clustering untuk Mengelompokkan Data Penayangan Film Bioskop Menggunakan Algoritma K-Means” yaitu tujuan utama dari penggunaan *Davies Bouldin Index* (DBI) adalah untuk mencapai tingkat optimal dari indeks tersebut sebagai hasil dari proses evaluasi klustering. DBI dijadikan tolak ukur kritis untuk mengukur efektivitas algoritma klustering, terutama dalam hal kemampuannya untuk membentuk *cluster-cluster* yang homogen, tetapi juga terpisah dengan baik satu sama lain. Dengan mencapai DBI yang optimal, penelitian atau implementasi algoritma klustering tidak hanya berupaya untuk meningkatkan tingkat keseragaman dan kemiripan di dalam setiap *cluster*, tetapi juga mengindikasikan upaya menuju struktur klaster yang paling sesuai dan representatif terhadap variasi data yang ada [6]. Penelitian yang dilakukan oleh Inas dan Ajeng Nur Afifah pada tahun 2023 yang berjudul “Data Mining Clustering Dalam Pengelompokan Buku Perpustakaan Menggunakan Algoritma K-Means” yaitu algoritma *K-Means* diakui sebagai sebuah teknik klasifikasi yang tidak hanya cepat tetapi juga sangat efisien dalam mengelompokkan data. Kecepatan algoritma ini terletak pada pendekatan iteratifnya yang mampu secara efisien mengelompokkan data ke dalam kelompok-kelompok berdasarkan kemiripan karakteristik. Keefisienan *K-Means* tercermin dalam kemampuannya untuk menangani jumlah data yang besar dengan cepat dan memberikan solusi klustering yang stabil [7]. Penelitian yang dilakukan oleh Mufidah Herviana dan Saleha Putri Delima pada tahun

2021 yang berjudul “Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Rawan Tanah Longsor Pada Provinsi Jawa Barat,” yaitu penentuan jumlah *cluster* optimal menggunakan Indeks *Davies-Bouldin* (DBI) merupakan pendekatan yang strategis dalam proses klustering, di mana tujuan utamanya adalah untuk menentukan jumlah klaster yang paling sesuai dan efektif untuk mewakili struktur intrinsik dari data yang diamati [8]. Penelitian yang dilakukan oleh Yeftha Christian pada tahun 2022 yang berjudul “Penerapan K-Means pada Segmentasi Pasar untuk Riset Pemasaran pada Startup Early Stage dengan Menggunakan CRISP-DM,” Implementasi algoritma *K-Means* untuk pengelompokan data merupakan langkah yang strategis dalam konteks analisis data dan pengambilan keputusan. Algoritma *K-Means*, sebagai metode klustering yang terkenal, diaplikasikan dengan tujuan untuk mengorganisir data ke dalam kelompok-kelompok yang homogen berdasarkan kemiripan karakteristik tertentu [9]. Penelitian yang dilakukan oleh Charles Zai pada tahun 2022 yang berjudul “Penggunaan Data Mining Memungkinkan Pengambilan Informasi yang Akurat dan Cepat” yaitu berisi penggunaan teknik data mining telah membuka peluang luas untuk pengambilan informasi yang tidak hanya akurat tetapi juga cepat dalam konteks analisis data. Dengan memanfaatkan algoritma dan metode data *mining* yang canggih, penelitian atau praktik analisis data dapat mengungkap pola-pola tersembunyi, hubungan korelasional, dan tren signifikan dari *volume* data yang kompleks.

2. TINJAUAN PUSTAKA

2.1. Peneliiian Terdahulu

Penelitian yang dilakukan oleh Charles Zai pada tahun 2022 yang berjudul “Penggunaan Data Mining Memungkinkan Pengambilan Informasi yang Akurat dan Cepat” yaitu berisi penggunaan teknik data mining telah membuka peluang luas untuk pengambilan informasi yang tidak hanya akurat tetapi juga cepat dalam konteks analisis data. Dengan memanfaatkan algoritma dan metode data *mining* yang canggih, penelitian atau praktik analisis data dapat mengungkap pola-pola tersembunyi, hubungan korelasional, dan tren signifikan dari *volume* data yang kompleks.

2.2. Data Mining

Data *mining* sebagai proses pemanfaatan metode statistik, matematika, kecerdasan buatan, dan machine learning untuk mengidentifikasi informasi yang bernilai dan menemukan pengetahuan yang tersembunyi dari berbagai basis data. Data yang ada digunakan untuk mengenali pola dan aturan umum yang dapat diterapkan pada klasifikasi data baru. Data *mining* memiliki peran utama dalam proses Penemuan Pengetahuan dalam Basis Data (*Knowledge Discovery in Database/KDD*), di mana algoritma-algoritma khusus digunakan untuk menggali data, membentuk

model, dan menemukan pola yang sebelumnya tidak diketahui [5].

Data mining adalah domain lintas disiplin yang menggabungkan metodologi dari berbagai bidang seperti pembelajaran mesin, pengenalan pola, analisis statistik, manajemen basis data, dan visualisasi data. Tujuan utamanya adalah untuk mengatasi tantangan yang terkait dengan pengambilan informasi dari database yang luas [10].

2.3. Algoritma K-Means

K-means merupakan metode pengelompokan data dimana penentuan cluster untuk setiap titik data bergantung pada tingkat keanggotaannya. Metode ini mengelompokkan dengan memisahkan data ke k daerah bagian yang terpisah, dimana k merupakan bilangan bulat positif. Implementasi algoritma K-Means sangat sederhana, cepat, dan mudah diadaptasi. Algoritma ini relative cepat dan efisien dalam mengolah data yang cukup besar, serta memiliki kemampuan yang besar dalam memproses data dengan waktu yang efisien (Nur Afifah & Nurdiyanto, 2023).

Algoritma K-means adalah teknik yang digunakan dalam bidang penambangan data. Ini berfungsi sebagai metode pengelompokan yang mencakup proses pemodelan tanpa pengawasan atau pembelajaran. Pendekatan untuk mengelompokkan data dalam algoritma ini dilakukan dengan cara yang dipartisi. Metodologi yang digunakan dalam algoritma K-Means melibatkan pengelompokan data menjadi beberapa bagian, dengan masing-masing kelompok memiliki karakteristik yang sama. Meskipun demikian, setiap kelompok juga menunjukkan sifat yang berbeda dibandingkan dengan kelompok lain. Tujuan utama dari algoritma ini adalah untuk meminimalkan perbedaan antara cluster sambil secara bersamaan memaksimalkan perbedaan antara titik data dari cluster yang berbeda [2].

2.4. Clustering

Clustering adalah proses pengelompokan data menjadi beberapa kelompok. Tujuan utama dari clustering data adalah untuk meminimalkan fungsi objektif dari proses clustering dan secara bersamaan memaksimalkan variasi antar kelompok [4].

Clustering adalah suatu proses di mana titik-titik data dikelompokkan ke dalam dua kelompok atau lebih. Tujuan utamanya adalah agar titik-titik data yang berada dalam kelompok yang sama memiliki kesamaan yang lebih tinggi daripada dengan kelompok yang berbeda, berdasarkan informasi yang tersedia dari titik-titik data tersebut [5].

2.5. Davies Bouldin Index

Indeks Davies-Bouldin (DBI) merupakan salah satu metode yang digunakan untuk mengevaluasi validitas atau jumlah cluster yang paling optimal dalam suatu metode pengelompokan. Indeks ini mengukur tingkat kedekatan data dengan titik pusat cluster dari klaster yang diikutinya. DBI mengukur

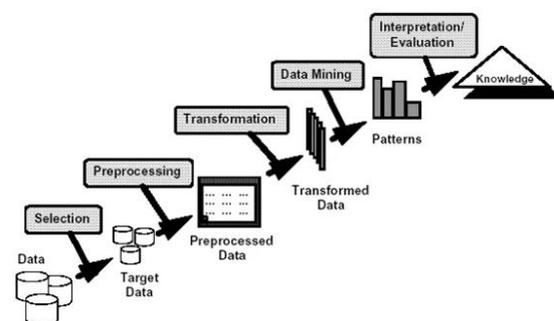
sejauh mana kluster-kluster tersebut saling berbeda dan sejauh mana setiap kluster memiliki kohesi yang tinggi [5].

Indeks Davies Bouldin (DBI) merupakan alat pengukur yang digunakan untuk mengevaluasi kinerja algoritma pengelompokan sebagai skema evaluasi. Ini memvalidasi seberapa baik pengelompokan dilakukan dengan menggunakan jumlah dan fitur pada dataset. DBI beroperasi dengan menghitung nilai rata-rata dari setiap elemen dalam kumpulan data. Jumlah klaster yang dipilih adalah jumlah klaster yang memiliki nilai DBI terendah, menunjukkan bahwa klaster tersebut memiliki kinerja terbaik menurut Davies Bouldin Index (DBI) [6].

Indeks Davies Bouldin (DBI) berdasarkan pada konsep kohesi dan separasi, di mana kohesi mencerminkan tingkat kedekatan data dengan pusat massa klaster yang sedang diamati, sementara separasi mengacu pada jarak antara pusat massa klaster [4].

3. METODE PENELITIAN

Metode yang digunakan dalam penelitian kali ini menggunakan metode perancangan Knowledge Discovery in Database (KDD), dengan menerapkan langkah demi langkah yang dilaksanakan secara berurutan untuk mencaoai tahap penyelesaian. Tahapan perancangan ini dapat dilihat pada gambar di bawah ini:



Gambar 1. Tahapan Proses KDD

Dalam penelitian ini, digunakan Metodologi Penemuan Pengetahuan dalam basis data (KDD) yang terstruktur dalam lima langkah, melibatkan pemilihan data, pre-pemrosesan atau pembersihan data, transformasi data, penambangan data untuk mencari pola atau informasi dan penafsiran atau evaluasi data.

Di bawah ini penjelasan terkait tahapan KDD sebagai berikut:

- a. *Data Selection*
Pemilihan data yang akan digunakan dalam proses data mining.
- b. *Pre-processing*
Tahap awal analisis data dimana data dibersihkan dari data yang missing.
- c. *Transformasi Data*
Mengubah Jenis data dari satu format ke format yang lain.

d. Data Mining

Tahap dimana algoritma atau teknik yang digunakan untuk mengolah data menggunakan algoritma *k-means clustering*.

e. Evaluation

Tahap terakhir hasil eksperimen di analisis untuk memahami data.

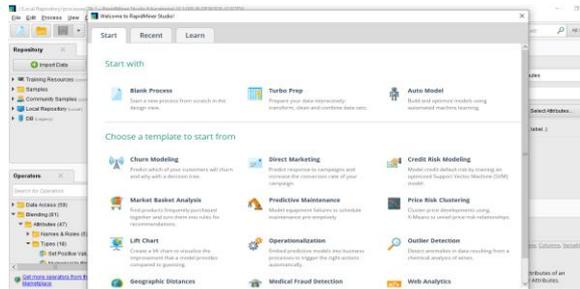
Pada tahapan yang pertama ini dilakukan data *selection* dimulai dengan pemilihan data yang akan digunakan dengan memilih atribut-atribut yang benar-benar berpengaruh dan berguna untuk menjawab pertanyaan atau tujuan analisis. Data yang digunakan sebanyak 12.544 *entry*. Tahap *Selection* bisa dilihat dari Tabel dibawah ini.

4. HASIL DAN PEMBAHASAN

Tabel 1. Data CSV PT Dika

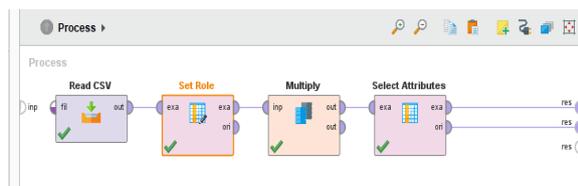
ID	USIA	PEKERJAAN	STATUS	KENDALA TRANSAKSI	KONTAK	SALDO
11781	21	murid	Lajang	No	telepon	4991.6
11782	42	kerah biru	telah menikah	No	seluler	5099.1
11783	42	teknisi	telah menikah	No	seluler	5076.2
11784	33	jasa	telah menikah	No	seluler	5099.1
11785	29	penganggur	Cerai	No	seluler	5017.5
11786	54	jasa	telah menikah	No	seluler	5228.1
11787	32	kerah biru	telah menikah	No	seluler	5099.1
11788	49	kerah biru	Cerai	No	seluler	5228.1
11789	44	teknisi	telah menikah	No	telepon	5191.2
11790	31	admin.	Cerai	No	telepon	5191.3
11791	45	teknisi	telah menikah	No	seluler	5228.1
11792	38	Pegawai	Cerai	No	telepon	5194.3
11793	31	jasa	telah menikah	No	seluler	5228.1
11794	49	jasa	telah menikah	No	seluler	5228.1

Langkah selanjutnya yaitu *pre-processing*. Langkah pertama membuka aplikasi *RapidMiner* untuk tampilan halaman awal *RapidMiner* bisa dilihat pada gambar dibawah ini.



Gambar 2. Tampilan Awal RapidMiner

Selanjutnya bisa menambahkan data dengan menggunakan atribut *Read CSV*, pada tahap selanjutnya yaitu transformasi data yang berarti melakukan suatu proses yang melibatkan perubahan data awal menjadi data khusus. Dalam data ini, terdapat 7 *attributes* dimana setelah melakukan proses *set role* menghasilkan 1 *attribute* khusus yaitu id. Di bawah ini adalah gambar pemodelan algoritma *k-means* pada *RapidMiner*.



Gambar 3. Pemodelan Algoritma K-Means

Dari pemodelan di atas kita dapat mengetahui hasil *cluster model* dari nilai k yang kita dapatkan. Menentukan nilai k dengan melihat hasil nilai *Davies Bouldin Index* (DBI) yang paling rendah. Berikut adalah tabel perbandingan nilai DBI.

Tabel 2. Tabel Perbandingan DBI

Nilai K	Nilai DBI
2	0,016
3	0,235
4	0,241
5	0,361
6	0,473
7	0,716
8	0,665
9	0,674
10	0,614
11	0,636
12	0,692
13	0,632
14	0,675
15	0,659
16	0,683
17	0,69
18	0,727
19	0,737
20	0,681

Pada proses *clustering* data tersebut, setelah melakukan perbandingan nilai K pada platform *RapidMiner*, ditemukan bahwa nilai *Davies-Bouldin Index* (DBI) yang optimal, mencerminkan kualitas klastering, diperoleh k memiliki nilai paling kecil yaitu 2. Hasil evaluasi menunjukkan pencapaian DBI sebesar 0,016. Nilai *Davies-Bouldin Index* yang rendah mencerminkan tingkat kompakness dan

separabilitas yang baik antar kluster, menandakan bahwa kluster yang dihasilkan oleh algoritma *clustering* memiliki karakteristik yang optimal.

Setelah melakukan pengecekan kembali menggunakan nilai $k = 2$ menghasilkan nilai DB dan *Cluster Model* yang bisa dilihat pada gambar berikut.

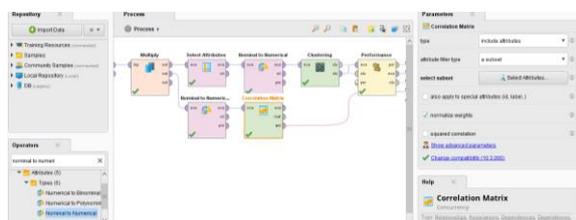


Gambar 4. Hasil Davies Bouldin



Gambar 5. Hasil Cluster Model

Berdasarkan hasil dari data diatas, bahwa proses dari *clustering* menghasilkan 2 kelompok (*cluster 0* dan *cluster 1*) dengan jumlah dalam setiap *cluster* berbeda. *Cluster 0* menghasilkan 10203 items, sedangkan pada *cluster 1* menghasilkan 2340 dengan total *number of items* adalah 12543. *Cluster 0* memiliki jumlah pelanggan yang lebih banyak, yaitu sekitar 70% dari total data. Hal ini menunjukkan bahwa sebagian besar adalah pelanggan setia. Pelanggan-pelanggan ini memiliki karakteristik yang cukup positif. Pada *cluster 1* memiliki jumlah pelanggan yang lebih sedikit, sekitar 30% dari total data. Hal ini menunjukkan bahwa penelitian ini memiliki potensi untuk mendapatkan pelanggan baru yang loyal.



Gambar 6. Correlation Matrix

Dalam operator *Correlation matrix*, terdapat penghilangan atribut yang memiliki nilai korelasi sebesar 0,014 terhadap atribut lainnya. Atribut yang memiliki nilai korelasi ini dianggap tidak memberikan kontribusi yang signifikan dalam pembentukan cluster. Proses seleksi ini dilakukan dengan tidak memilih atribut tersebut dalam operator *selection*.

Pada proses yang melibatkan homogenitas dan separabilitas kluster, karena atribut yang memiliki korelasi tinggi cenderung mencerminkan pola-pola yang serupa di dalam data.

5. KESIMPULAN

Hasil analisis menunjukkan bahwa kluster 2 memberikan nilai terbaik berdasarkan Davies Boulding Index (DBI) dengan mencapai nilai 0,016. Dalam pengelompokan dua kluster yang telah ditentukan, kluster 0 memiliki 10.203 item, sementara kluster 1 memiliki 2.340 item. Namun, perlu dicatat bahwa hasil dari operator correlation matrix menghasilkan DBI yang sedikit lebih rendah, yaitu 0,014. Oleh karena itu, saran yang dapat diajukan adalah untuk penelitian selanjutnya, sebaiknya peneliti mempertimbangkan penggunaan metode selain K-Means Clustering dalam mengelompokkan jumlah penduduk. Hal ini dapat membuka peluang untuk mencari metode clustering yang lebih sesuai atau dapat memberikan hasil yang lebih optimal dalam konteks penelitian tersebut. Kesimpulan ini dapat membantu meningkatkan keakuratan dan ketepatan hasil pengelompokan populasi, serta memperluas pemahaman terhadap karakteristik masyarakat yang lebih kompleks.

DAFTAR PUSTAKA

- [1] E. M. Fitri, R. R. Suryono, and A. Wantoro, "Klasterisasi Data Penjualan Berdasarkan Wilayah Menggunakan Metode K-Means Pada Pt Xyz," *J. Komputasi*, vol. 11, no. 2, pp. 157–168, 2023, doi: 10.23960/komputasi.v11i2.12582.
- [2] F. N. Dhewayani, D. Amelia, D. N. Alifah, B. N. Sari, and M. Jajuli, "Implementasi K-Means Clustering untuk Pengelompokkan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM," *J. Teknol. dan Inf.*, vol. 12, no. 1, pp. 64–77, 2022, doi: 10.34010/jati.v12i1.6674.
- [3] M. Syaefudulloh, A. Faqih, and F. M. Basysyar, "Clustering Kelompok Belajar Siswa Berdasarkan Hasil Ujian Sekolah Menggunakan Algoritma K-Means," *J. Sist. Inf. dan Manaj.*, vol. Volume 10, no. 1, pp. 195–199, 2022.
- [4] N. Suarna and Y. A. Wijaya, "Analisa Penerapan Metode Clustering K-Means Untuk Pengelompokan Data Trransaksi Konsumen (Studi Kasus : Cv . Mitra Indexindo Pratama)," vol. 7, no. 2, pp. 1322–1328, 2023.
- [5] A. Permadi and Y. A. Wijaya, "Pengelompokkan Dataset Bus Menggunakan Algoritma K-Means," *Informatics Educ. ...*, vol. 7, no. 2, pp. 138–152, 2023, [Online]. Available: <http://ejournal-binainsani.ac.id/index.php/ITBI/article/view/22590Ahttps://ejournal-binainsani.ac.id/index.php/ITBI/article/download/2259/1557>

- [6] M. N. Dayat, N. Suarna, and Y. A. Wijaya, "Analisa Clustering untuk Mengelompokkan Data Penayangan Film Bioskop Menggunakan Algoritma K-Means," *Intern. (Information Syst. Journal)*, vol. 6, no. 1, pp. 68–78, 2023, [Online]. Available: <https://jurnal.masoemiversity.ac.id/index.php/internal/article/view/686>
- [7] I. A. Nur Afifah and H. Nurdiyanto, "Data Mining Clustering Dalam Pengelompokan Buku Perpustakaan Menggunakan Algoritma K-Means," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 8, no. 3, pp. 802–814, 2023, doi: 10.29100/jupi.v8i3.3891.
- [8] M. Herviany, S. Putri Delima, T. Nurhidayah, and K. Kasini, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokkan Daerah Rawan Tanah Longsor Pada Provinsi Jawa Barat," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 34–40, 2021, doi: 10.57152/malcom.v1i1.60.
- [9] Y. Christian and K. O. Y. R. Qi, "Penerapan K-Means pada Segmentasi Pasar untuk Riset Pemasaran pada Startup Early Stage dengan Menggunakan CRISP-DM," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, p. 966, 2022, doi: 10.30865/jurikom.v9i4.4486.
- [10] F. S. Amalia, S. Setiawansyah, and ..., "Analisis Data Penjualan Handphone Dan Elektronik Menggunakan Algoritma Apriori (Studi Kasus: Cv Rey Gasendra)," ... *J. Telemat. ...*, vol. 2, no. 1, pp. 1–6, 2021.