PENINGKATAN MODEL SEGMENTASI PENGGUNA TERMINAL TIPE B SUMBER KABUPATEN CIREBON DALAM PERBAIKAN SARANA DAN PRASARANA DENGAN ALGORITMA K-MEANS

Agesty Kusmiyaty, Rudi Kurniawan, Yudhistira Arie Wijaya, Umi Hayati
Teknik Informatika, STMIK IKMI Cirebon
Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat, Indonesia
agestykusmiyaty2301@gmail.com

ABSTRAK

Terminal Sumber, sebuah terminal tipe B di Kabupaten Cirebon, menghadapi tantangan dalam memenuhi kebutuhan pengguna akibat pengelolaan sarana dan prasarana yang belum optimal. Tingkat kepuasan pengguna terhadap fasilitas, kebersihan, kualitas layanan, dan infrastruktur sering kali beragam, mencerminkan kebutuhan perbaikan yang terarah. Penelitian ini bertujuan untuk menganalisis kepuasan pengguna dengan memanfaatkan algoritma K-Means Clustering. Data kepuasan pengguna dikumpulkan melalui survei langsung, menghasilkan dua klaster berdasarkan nilai Davies Bouldin Index (DBI) optimal sebesar 0,547. Klaster 0 merepresentasikan pengguna dengan tingkat kepuasan lebih rendah, sedangkan klaster 1 memiliki tingkat kepuasan lebih tinggi. Hasil ini menunjukkan bahwa segmentasi pengguna dapat membantu pengelola terminal merancang strategi perbaikan layanan yang lebih spesifik dan efektif. Dengan demikian, penelitian ini berkontribusi dalam penerapan metode clustering untuk meningkatkan kualitas layanan transportasi umum, baik secara praktis maupun akademis.

Kata kunci : *Segmentasi Kepuasan Pengguna, Terminal, Perbaikan Sarana dan Prasarana, Algoritma K-Means, Nilai DBI (Davies Bouldin Index)*

1. PENDAHULUAN

Perkembangan pesat di bidang informatika telah memberikan dampak signifikan pada berbagai aspek kehidupan, termasuk teknologi, bisnis, dan pelayanan publik. Salah satu teknologi yang memiliki peran penting dalam analisis data besar adalah algoritma clustering [7]. Algoritma ini, khususnya K-Means, telah terbukti efektif dalam mengelompokkan data untuk berbagai tujuan, seperti analisis data pendidikan [8]. Salah satu contoh penerapannya adalah dalam segmentasi pengguna untuk meningkatkan efisiensi layanan publik, seperti transportasi umum, sehingga menunjukkan bahwa pendekatan berbasis clustering dapat memberikan nilai tambah yang signifikan dalam pengelolaan data pengguna dan segmentasi pelanggan [6].

Dalam konteks transportasi umum, penerapan algoritma K-Means menjadi semakin relevan untuk mengidentifikasi pola penggunaan layanan transportasi berdasarkan data pengguna yang beragam [2]. Terminal Tipe B Sumber di Kabupaten Cirebon merupakan salah satu fasilitas transportasi yang memerlukan strategi pengelolaan yang efektif agar dapat memenuhi kebutuhan pengguna secara optimal. Oleh karena itu, penerapan algoritma K-Means dalam segmentasi pengguna terminal ini dapat menjadi solusi dalam mengidentifikasi dan mengelompokkan berdasarkan pola serta preferensi mereka, sehingga perbaikan sarana dan prasarana dapat lebih tepat sasaran [15].

Penelitian sebelumnya menunjukkan bahwa segmentasi pengguna terminal transportasi sering kali belum memanfaatkan data secara optimal, sehingga perbaikan yang dilakukan masih bersifat umum dan kurang spesifik [2]. Hal ini menyebabkan efisiensi pengelolaan sarana dan prasarana terminal menurun, dan kepuasan pengguna pun tidak tercapai. Dalam konteks terminal Tipe B Sumber di Kabupaten Cirebon, pendekatan segmentasi yang lebih mendalam menggunakan algoritma K-Means menjadi relevan untuk diterapkan, karena algoritma ini mampu mengelompokkan data dalam jumlah besar dengan presisi tinggi [15].

Penelitian terkait penerapan algoritma K-Means telah dilakukan dalam berbagai konteks. Menunjukkan bahwa algoritma K-Means berhasil digunakan untuk menganalisis kepuasan pengguna layanan e-commerce [14]. Memanfaatkan algoritma ini untuk segmentasi pelanggan salon kecantikan, yang menunjukkan hasil signifikan dalam memahami kebutuhan spesifik pelanggan [9]. Namun, kajian terkait penerapan algoritma ini dalam konteks fasilitas transportasi, seperti terminal, masih sangat terbatas [5].

Tujuan utama penelitian ini adalah untuk mengembangkan model segmentasi pengguna Terminal Tipe B Sumber Kabupaten Cirebon menggunakan algoritma K-Means. Dengan mengidentifikasi kelompok pengguna berdasarkan pola dan preferensi mereka, penelitian ini diharapkan dapat memberikan kontribusi nyata dalam meningkatkan pengelolaan fasilitas publik secara efisien. Selain itu, hasil penelitian ini dapat menjadi pedoman bagi pengelola terminal dalam merancang strategi perbaikan sarana dan prasarana yang lebih sesuai dengan kebutuhan [10].

Penelitian ini akan menggunakan pendekatan kuantitatif dengan penerapan algoritma K-Means untuk segmentasi data pengguna terminal. Metode ini dipilih karena kemampuannya dalam mengelompokkan data secara efektif berdasarkan

karakteristik pengguna. Analisis data dilakukan dengan memanfaatkan perangkat lunak komputasi yang relevan, serta evaluasi hasil clustering menggunakan indeks evaluasi seperti Davies-Bouldin Index untuk memastikan akurasi hasil segmentasi [4].

2. TINJAUAN PUSTAKA

2.1. Algoritma K-Means

Algoritma K-Means adalah salah satu metode clustering yang populer dan sering digunakan dalam pengelompokan data berdasarkan kedekatan atribut. Algoritma ini sangat efektif dalam menganalisis data kinerja dosen untuk menghasilkan pengelompokan yang relevan dengan pola yang ditemukan [2]. Menunjukkan bagaimana K-Means dapat diterapkan untuk memahami kebutuhan pelanggan, terutama di sektor jasa, yang memberikan wawasan untuk meningkatkan layanan sesuai preferensi pelanggan [9].

Algoritma ini juga efektif dalam sektor transportasi untuk meningkatkan layanan berbasis aplikasi [12]. Pada penelitian ini, algoritma K-Means digunakan untuk memisahkan pengguna berdasarkan tingkat kepuasan mereka terhadap layanan terminal, menciptakan segmentasi yang dapat diandalkan.

2.2. Davies Bouldin Index (DBI)

DBI adalah metrik yang digunakan untuk mengevaluasi kualitas cluster. Semakin rendah nilai DBI, semakin baik kualitas cluster yang terbentuk karena memiliki jarak intra yang kecil dan jarak inter yang besar. DBI membantu dalam menentukan jumlah cluster optimal, terutama dalam konteks transportasi public[15].

Sebagai contoh, bahwa DBI dapat digunakan secara efektif dalam menentukan segmentasi pelanggan pada e-commerce [14]. Dalam studi ini, pendekatan serupa diterapkan untuk mengevaluasi kualitas cluster pengguna terminal.

2.3. Segmentasi Pengguna

Segmentasi pengguna memungkinkan pemahaman yang lebih baik terhadap preferensi dan kebutuhan kelompok tertentu. Segmentasi ini telah diterapkan dalam berbagai bidang, seperti pendidikan [13] dan transportasi [4]. Segmentasi berbasis data memberikan fokus pada kebutuhan spesifik, sehingga layanan dapat disesuaikan dengan lebih baik.

Penelitian ini memanfaatkan segmentasi untuk membedakan pengguna dengan kepuasan rendah dari mereka yang memiliki kepuasan tinggi. Hasilnya memberikan rekomendasi yang lebih relevan kepada pengelola terminal.

2.4. Integrasi Teknologi Dalam Transportasi

Pentingnya integrasi teknologi dalam layanan transportasi, seperti sistem pelacakan real-time dan aplikasi berbasis mobile[4]. Teknologi ini memberikan nilai tambah bagi pengguna muda yang lebih terhubung dengan perangkat digital.

Teknologi berbasis data dapat digunakan untuk memantau dan meningkatkan kualitas layanan secara real-time, meningkatkan kepercayaan dan kenyamanan pengguna[8].

2.5. Penggunaan Data Untuk Efisiensi Operasional

Data memainkan peran penting dalam meningkatkan efisiensi operasional. Pentingnya data dalam mengelola sumber daya di sektor e-commerce, seperti penjadwalan dan alokasi[14]. Pendekatan ini dapat diterapkan di transportasi publik untuk memastikan pengelolaan fasilitas dan jadwal yang optimal berdasarkan kebutuhan pengguna.

2.6. Pengelolaan Infrastruktur Publik

Pengelolaan infrastruktur publik berbasis data dapat mengidentifikasi kekurangan dalam fasilitas. Penelitian ini memberikan rekomendasi spesifik, seperti perbaikan ruang tunggu, kebersihan, dan aksesibilitas, yang relevan bagi pengguna dengan kepuasan rendah[5].

2.7. Analisis Data Demografis

Analisis data demografis, seperti usia dan pendapatan, sangat membantu dalam memahami kebutuhan spesifik dari berbagai kelompok pengguna. Data demografis dapat digunakan untuk menentukan prioritas peningkatan fasilitas, seperti ruang tunggu yang nyaman untuk lansia atau informasi berbasis aplikasi untuk pengguna muda [13].

2.8. Evaluasi Kinerja Layanan Publik

Evaluasi kinerja layanan publik dengan metode clustering memberikan wawasan berharga. Evaluasi ini dapat menjadi dasar untuk peningkatan layanan yang lebih terfokus. Hasil ini relevan dalam konteks terminal yang membutuhkan pembaruan layanan dinamis[16].

2.9. Pemanfaatan Model Clustering Lain

Selain K-Means, model clustering lain seperti DBSCAN dapat digunakan untuk membandingkan hasil segmentasi. DBSCAN cocok untuk pola data yang lebih kompleks, memberikan fleksibilitas dalam evaluasi segmentasi pengguna. [4]

2.10. Keberlanjutan Penelitian Transportasi Publik

Penelitian berkelanjutan dalam transportasi publik sangat penting untuk memastikan adaptasi terhadap perubahan kebutuhan pengguna. Model berbasis data dapat menghasilkan wawasan yang relevan untuk analisis mendalam di masa depan, membantu pengelola terminal dalam merencanakan peningkatan fasilitas. [10]

3. METODE PENELITIAN

3.1. Sumber Data

Pada penelitian ini, sumber data yang digunakan terdiri dari data primer dan data sekunder. Data primer diperoleh melalui survei kepuasan pengguna Terminal Sumber, yang dilakukan untuk mendapatkan informasi langsung mengenai pengalaman pengguna dalam menggunakan fasilitas terminal. Metode pengumpulan data ini dipilih karena dapat memberikan informasi yang relevan dan terkini. Selain itu, data sekunder digunakan berupa literatur dan artikel ilmiah terkait teknik K-Means dalam pengelompokan data dan aplikasi teknik ini pada analisis kepuasan pengguna, yang diperoleh dari sumber terpercaya seperti jurnal ilmiah, laporan penelitian, dan basis data publik [2]. Kualitas data dievaluasi berdasarkan validitas, reliabilitas, dan akurasi, di mana data survei dipastikan mencerminkan populasi pengguna terminal, dan data sekunder berasal dari sumber yang dapat diandalkan. Ketersediaan data dipastikan dengan memeriksa izin dan lisensi yang diperlukan untuk mengakses data sekunder, serta memastikan data primer dapat diperoleh dengan mudah [7]. Semua data yang digunakan dalam penelitian ini relevan dengan tujuan penelitian, yaitu untuk menganalisis kepuasan pengguna terminal dan mengelompokkan data berdasarkan karakteristik yang dapat mendukung perbaikan layanan terminal.

3.2. Populasi dan Sampel

Populasi dalam penelitian ini adalah seluruh pengguna Terminal Sumber yang menggunakan berbagai fasilitas yang tersedia di terminal, seperti transportasi, ruang tunggu, dan pelayanan lainnya. Metode sampling yang digunakan adalah random sampling, di mana sampel dipilih secara acak untuk memastikan setiap individu dalam populasi memiliki kesempatan yang sama untuk dipilih, sehingga dapat menghasilkan sampel yang representatif. Ukuran sampel ditentukan dengan mempertimbangkan tingkat kepercayaan yang diinginkan dan margin of error yang dapat diterima, sehingga sampel yang digunakan cukup besar untuk mewakili populasi secara keseluruhan [8]. Pengambilan sampel dilakukan dengan mengidentifikasi responden secara acak dari pengguna terminal yang ada pada saat survei, memastikan bahwa sampel mencakup variasi dalam karakteristik pengguna, seperti usia, jenis kelamin, dan jenis layanan yang digunakan, untuk mendapatkan gambaran yang lebih komprehensif mengenai kepuasan pengguna terminal.

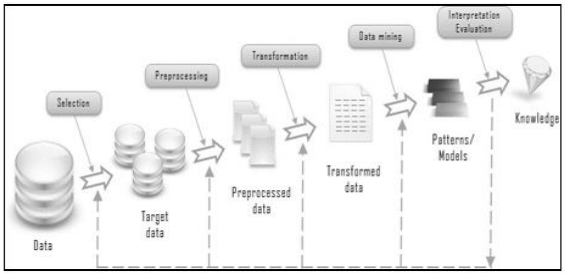
3.3. Teknik Pengumpulan Data

Setelah sampel ditentukan, pengumpulan data dilakukan melalui survei kepuasan pengguna Terminal Sumber. Survei ini disusun dengan menggunakan kuesioner yang berisi pertanyaan terkait pengalaman dan kepuasan pengguna terhadap fasilitas dan layanan yang disediakan di terminal. Kuesioner dirancang untuk mengumpulkan data kuantitatif yang dapat dianalisis menggunakan teknik K-Means clustering. Proses pengumpulan data dilakukan secara langsung di lokasi terminal dengan meminta responden yang terpilih untuk mengisi kuesioner secara sukarela [14]. Survei ini melibatkan pertanyaan yang mencakup berbagai aspek layanan, seperti kenyamanan fasilitas, waktu tunggu, kebersihan, dan pelayanan petugas, untuk mendapatkan gambaran menyeluruh tentang kepuasan pengguna. Data yang terkumpul kemudian dianalisis untuk mengidentifikasi pola dan tren dalam kepuasan pengguna berdasarkan klaster yang terbentuk. Teknik pengumpulan data ini dipilih karena memungkinkan untuk mendapatkan informasi yang akurat dan relevan langsung dari pengguna yang berinteraksi dengan layanan terminal.

.

3.4. Teknik Analisis Data

Setelah data terkumpul melalui survei kepuasan pengguna, analisis data dilakukan menggunakan pendekatan Knowledge Discovery in Databases (KDD), yang melibatkan serangkaian proses untuk mengekstraksi pengetahuan dari data yang besar dan kompleks. Dalam penelitian ini, KDD digunakan untuk memproses data survei, dimulai dengan seleksi, pembersihan, dan transformasi data sebelum melakukan pemodelan dengan algoritma K-Means clustering, yang bertujuan untuk mengelompokkan responden berdasarkan tingkat kepuasan mereka terhadap berbagai aspek layanan di Terminal Sumber [4]. Hasil analisis ini kemudian diinterpretasikan dengan membandingkan tingkat kepuasan antar klaster untuk mengidentifikasi faktor-faktor yang mempengaruhi kepuasan pengguna. Temuan yang diperoleh akan memberikan wawasan yang berguna dalam merumuskan rekomendasi perbaikan pelayanan, dan jika hasil ini digunakan untuk menggeneralisasi populasi pengguna terminal secara keseluruhan, perlu disampaikan batasan serta implikasi dari generalisasi tersebut, termasuk potensi bias dan keterbatasan dalam ukuran sampel yang digunakan [5]. Dengan demikian, pendekatan KDD ini memberikan dasar yang kuat dalam merancang kebijakan dan strategi untuk meningkatkan kualitas layanan dan kepuasan pengguna terminal.

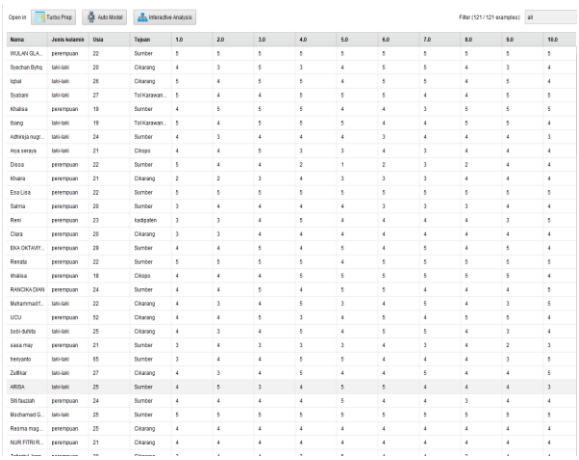


Gambar 1. Alur tahapan *Knowledge Discovery in Databases (KDD)*
Sumber: Jurnal Ilmiah Teknologi Informasi Asia (2022)

4. HASIL DAN PEMBAHASAN

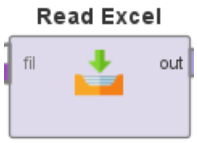
4.1. Sumber Data

Penelitian ini memanfaatkan data yang dikumpulkan melalui survei kepuasan pengguna Terminal Sumber. Data tersebut menggambarkan persepsi dan penilaian pengguna terhadap sarana dan prasarana yang ada di Terminal Sumber. Informasi yang diperoleh dari survei, yang mencakup 20 atribut dan 121 baris data, divisualisasikan pada gambar berikut:



Gambar 2. Data survei pengguna terminal Sumber

Data ini akan diolah menggunakan Rapid Miner untuk menganalisis dan memvisualisasikan hasil secara efektif.



Gambar 3. Operator read excel

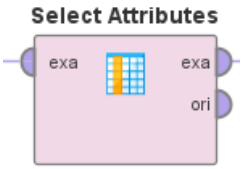
Menampilkan penggunaan operator Read Excel untuk memuat data survei ke dalam RapidMiner. Operator Read Excel berfungsi untuk mengimpor data dari file Excel (.xlsx atau .xls) ke dalam proses analisis.

4.2. Atribut Yang Diseleksi

Pemilihan atribut didasarkan pada relevansi dan kontribusinya terhadap analisis klasterisasi data survey pengguna Terminal Sumber. Atribut 'Nama' berfungsi sebagai pengenal unik untuk setiap material dan dipertahankan sebagai label. Atribut '1' hingga '15' ditetapkan sebagai atribut target yang akan di klasterisasi. Atribut-atribut lainnya bertindak sebagai atribut pendukung yang memberikan informasi tambahan mengenai survey. Atribut 'Saran' dieksklusikan karena berisi informasi subjektif dan opsional. Tabel 1 menampilkan atribut yang dipilih dalam RapidMiner:

Tabel 1. Atribut yang diseleksi

attributes	sellected attributes
No	1-15
	jenis kelamin
	jurusan
	nama
	usia



Gambar 4. Operator select attributes

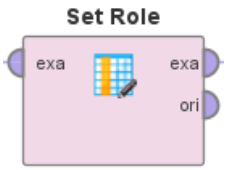
Operator pada gambar diatas adalah Select Attributes Fungsi dari operator ini adalah untuk memilih atribut tertentu dari dataset yang digunakan dalam proses analisis data.

4.3. Labeling

Tabel 2 menampilkan penetapan atribut 'Nama' sebagai label. Proses pelabelan ini bertujuan untuk menyederhanakan analisis dan interpretasi hasil klasterisasi. Dengan menetapkan 'Nama' sebagai label, setiap cluster yang terbentuk dapat diidentifikasi dan diinterpretasikan berdasarkan karakteristik yang tergabung dalam cluster tersebut. Hal ini memungkinkan untuk memahami pola dan karakteristik yang mendasari pengelompokan data secara lebih mudah dan komprehensif.

Tabel 2. Atribut yang diseleksi

Atribute name	Target Role
Nama	Label



Gambar 5. Operator set role

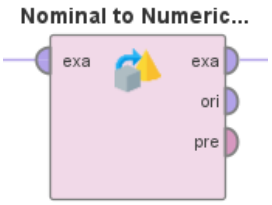
Operator Set Role di RapidMiner digunakan untuk mengubah peran (role) dari atribut dalam dataset. Peran atribut sangat penting dalam menentukan bagaimana atribut tersebut digunakan dalam proses analisis. Fungsi utama dari operator ini adalah untuk mengubah Nama sebagai label.

4.4. Merubah Atribut Nominal Ke Numerical

Pengkodean Unique Integer, seperti yang ditunjukkan pada Tabel 3, diaplikasikan pada atribut kategorikal 'Jenis Kelamin' dan 'Tujuan'. Tujuan dari transformasi ini adalah untuk mengonversi data atribut menjadi representasi numerik yang kompatibel dengan algoritma K-Means.

Tabel 3. Merubah atribut nominal ke numerik

attributes	sellected attributes
nama	jenis kelamin
	jurusan



Gambar 6. Nominal to numerical

Opeartor Nominal to Numerical merupakan Transformasi ini memungkinkan algoritma K-Means memproses data dengan lebih efisien. Atribut kategorikal, seperti 'Jenis Kelamin' dan untuk mengubah data kategorikal menjadi numerik.

4.5. Mengganti Nilai Yang Hilang

Penanganan nilai yang hilang, seperti yang ditunjukkan pada Tabel 4, dilakukan dengan mengganti data yang hilang pada atribut yang digunakan untuk analisis menggunakan nilai rata-rata atribut. Pendekatan ini dijustifikasi oleh proporsi data yang hilang yang relatif kecil, sehingga diasumsikan tidak akan mempengaruhi hasil analisis secara substansial.

Tabel 4. Mengganti nilai yang hilang

attributes	sellected attributes
jenis kelamin	1
jurusan	2
nama	3
usia	4
	5
	6
	7
	8
	9
	10
	11
	12
	13
	14
	15

Gambar 7. Operator respalce missing value

Operator Replace Missing Values pada RapidMiner digunakan untuk menangani nilai yang hilang (missing values) dalam dataset. Nilai yang hilang seringkali menjadi masalah dalam analisis data, karena banyak algoritma tidak dapat bekerja dengan baik jika ada nilai yang kosong.

4.6. Normalisasi Data

Seperti yang ditunjukkan pada Tabel 6, data dinormalisasi menggunakan Normalize untuk mengatasi perbedaan skala antar atribut. Hal ini dilakukan untuk menghindari bias dalam proses pengelompokan (clustering) yang dapat muncul jika atribut dengan skala lebih besar mendominasi perhitungan jarak. Normalisasi memastikan bahwa semua atribut memiliki bobot yang sebanding dalam analisis.

Tabel 5. Normalisasi data

attributes	sellected attributes
jenis kelamin	1-15
tujuan	
usia	

Gambar 8. Operator normalize

Operator Normalize penting untuk meningkatkan performa model dan memastikan analisis yang lebih akurat pada dataset dengan atribut yang memiliki skala berbeda.

4.7. Penentuan Nilai K Optimal

Dalam penelitian ini, algoritma K-Means diimplementasikan untuk melakukan pengelompokan (clustering) terhadap data stok bahan manufaktur. Tahap awal yang krusial adalah menentukan jumlah cluster (K) yang optimal. Untuk mencapai tujuan ini, Davies-Bouldin Index (DBI) dipilih sebagai metrik evaluasi.DBI mengukur kualitas cluster dengan mempertimbangkan rasio antara jarak antar cluster dan jarak dalam cluster. Nilai DBI yang lebih rendah mengindikasikan kualitas cluster yang lebih baik, karena menunjukkan cluster yang lebih padat dan terpisah dengan baik.Secara matematis, DBI didefinisikan sebagai rata-rata similaritas maksimum antara setiap cluster dengan cluster tetangganya. Similaritas dihitung berdasarkan perbandingan jarak rata-rata dalam cluster dengan jarak antar pusat cluster. Tabel 6 menampilkan deretan nilai DBI terbaik dari hasil experiment

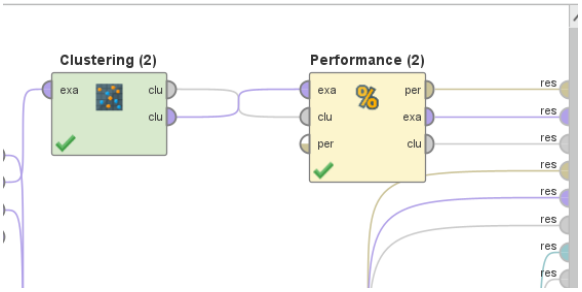
Tabel 6. Menampilkan nilai DBI

Interatior	Cluster	DBI
1	2	0.547
4	9	1.363
6	14	1.330
2	4	0.903
5	12	1.295
8	19	1.275
9	21	1.272
7	16	1.279
3	7	1.157
10	24	1.166
16	38	1.145
15	36	1.017
12	28	1.224
14	33	1.183
17	40	1.082
18	43	1.087
13	31	1.191
21	50	1.001
20	48	0.959
19	45	1.029
11	26	1.313

Pada Tabel 6 tersebut, diperlihatkan bahwa nilai K optimal yang diperoleh adalah K = 2, dengan nilai Indeks Davies-Bouldin (DBI) sebesar 0,547. Nilai ini menunjukkan bahwa pemilihan K = 2 memberikan hasil klasterisasi yang paling efektif berdasarkan kriteria DBI, yang merupakan ukuran kualitas klaster dengan mempertimbangkan tingkat kepadatan dan keterpisahan antar klaster.

4.8. Melakukan K Means Cluster

Algoritma K-Means diterapkan pada data survei pengguna terminal sumber yang telah melalui proses normalisasi, menggunakan nilai K optimal yang telah ditentukan sebelumnya. Hasil klasterisasi mengindikasikan bahwa data terbagi menjadi dua klaster utama. Di tampilkan pada Gambar 9 Operator yang digunakan dalam RapidMiner



Gambar 9. Operator K-Means clustering

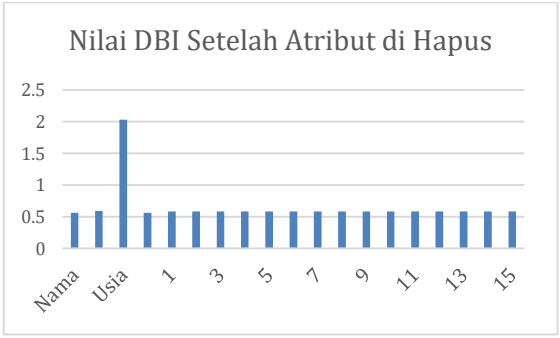
K-means clustering adalah algoritma unsupervised learning yang mengorganisir data ke dalam beberapa kelompok berdasarkan kemiripan atribut, melalui proses iterasi hingga stabil.

4.9. Mencari Atribut Berpengaruh

Setelah menentukan nilai K optimal, dilakukan analisis untuk mengidentifikasi atribut yang paling berpengaruh terhadap pembentukan klaster. Analisis ini dilakukan dengan menghitung nilai Davies-Bouldin Index (DBI) setelah menghapus satu per satu atribut dari dataset. Atribut yang jika dihapus menyebabkan penurunan nilai DBI yang paling besar dianggap sebagai atribut yang paling berpengaruh. Tabel 7 menunjukkan nilai DBI setelah menghapus masing-masing atribut.

Tabel 7. Nilai DBI dari masing-masing atribut yang hapus

Nama Atribut	DBI dari K 2
Nama	0,563
Jenis kelamin	0,59
Usia	2,033
Tujuan	0,563
1	0,585
2	0,585
3	0,585
4	0,585
5	0,585
6	0,585
7	0,585
8	0,585
9	0,585
10	0,585
11	0,585
12	0,585
13	0,585
14	0,585
15	0,585



Gambar 10. Nilai DBI setelah atribut di hapus

Pada Gambar 9 dapat disimpulkan bahwa, atribut 'Usia' diidentifikasi sebagai faktor dengan pengaruh terbesar dalam pembentukan klaster. Hal ini terlihat dari kenaikan nilai Davies-Bouldin Index (DBI) yang paling tinggi, yaitu mencapai 2,033, ketika atribut 'Usia' dikeluarkan dari dataset.

4.10. Karakteristik Dari Masing-Masing Cluster

Karakteristik setiap cluster dijelaskan berdasarkan nilai rata-rata dan standar deviasi atribut-atributnya. Label atau nama yang representatif diberikan kepada setiap cluster berdasarkan karakteristik yang dominan. Untuk memperjelas karakteristik masing-masing cluster, disajikan Gambar 4.10 yang merangkum nilai rata-rata dan standar deviasi atribut-atribut untuk setiap cluster.

Tabel 8 Karakteristik masing-masing cluster

Atribut	Cluster_0	Cluster_1
1.0	3.806	3.469
2.0	3.806	3.735
3.0	4.194	3.755
4.0	4.139	3.857
5.0	4.014	4.041
6.0	4.111	3.939
7.0	4.125	3.878
8.0	3.903	3.776
9.0	4,083	3.796
10.0	4.097	3.918
11.0	3.931	4.0
12.0	3.861	3.816
13.0	3.833	4.020
14.0	3.972	3.878
15.0	3.903	3.796
Jenis Kelamin	0.431	0.551
Jurusan	1,764	3.041
Usia	25.069	48.694

Tabel 8 menunjukkan bahwa klaster_0 didominasi oleh pengguna dengan tingkat kepuasan di bawah rata-rata dan usia di atas 40 tahun. Sementara itu, klaster_1 didominasi oleh pengguna dengan tingkat kepuasan di atas rata-rata dan usia rata-rata 25 tahun.

4.11. Interpretasi Hasil Dan Keterkaitan Dengan Literatur

Penelitian ini bertujuan untuk menentukan jumlah klaster (K) optimal menggunakan algoritma K-Means. Berdasarkan hasil yang ditunjukkan pada Gambar 4.7, nilai K optimal yang diperoleh adalah 2, dengan nilai Davies-Bouldin Index (DBI) sebesar 0,547. Nilai DBI yang rendah ini menunjukkan kualitas klaster yang baik, dengan klaster yang padat dan terpisah secara signifikan. Hasil ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa penggunaan DBI dapat memberikan evaluasi objektif dalam menentukan jumlah cluster optimal [16].

Deskripsi dari masing-masing klaster memberikan gambaran karakteristik pengguna berdasarkan hasil klasterisasi. Gambar 4.10 menunjukkan bahwa klaster_0 terdiri dari pengguna dengan tingkat kepuasan yang berada di bawah rata-rata dan usia di atas 40 tahun. Sebaliknya, klaster_1 mencakup pengguna dengan tingkat kepuasan yang lebih tinggi dari rata-rata dan usia sekitar 25 tahun. Perbedaan ini mengindikasikan adanya segmentasi yang signifikan berdasarkan usia dan tingkat kepuasan pengguna. Faktor demografis, seperti usia, dapat mempengaruhi tingkat kepuasan pengguna dalam konteks yang berbeda[17].

Atribut yang paling berpengaruh terhadap nilai DBI diidentifikasi melalui analisis penghapusan atribut satu per satu dari dataset. Berdasarkan Gambar 4.9, atribut 'Usia' ditemukan memiliki dampak paling signifikan terhadap nilai DBI. Penghapusan atribut ini mengakibatkan kenaikan DBI hingga mencapai 2,033, yang menunjukkan bahwa atribut 'Usia' memberikan kontribusi besar dalam membentuk klaster yang lebih terstruktur. Variabel demografis, seperti usia, memainkan peran penting dalam analisis segmentasi pelanggan[18].

5. KESIMPULAN DAN SARAN

Hasil penelitian menunjukkan bahwa pengelompokan kepuasan pengguna Terminal Sumber dengan algoritma K-Means menghasilkan dua cluster terbaik dengan nilai Davies-Bouldin Index (DBI) sebesar 0,547, di mana cluster 1 menunjukkan tingkat kepuasan lebih tinggi dibandingkan cluster 0. Analisis atribut mengungkapkan bahwa usia memiliki pengaruh signifikan dalam pembentukan cluster, dengan cluster 1 didominasi oleh pengguna yang lebih muda dan lebih puas, sedangkan cluster 0 terdiri dari pengguna yang lebih tua dengan tingkat kepuasan lebih rendah. Penemuan ini memberikan wawasan praktis bagi pengelola terminal untuk mengidentifikasi kebutuhan spesifik dan merancang perbaikan layanan yang lebih terarah, seperti peningkatan fasilitas untuk pengguna yang lebih tua dan layanan tambahan bagi kelompok muda. Untuk penelitian lanjutan, disarankan memperluas cakupan dataset, menambahkan atribut yang relevan seperti keamanan dan kebersihan, serta menggunakan algoritma clustering alternatif untuk validasi hasil. Selain itu, pengelola terminal diharapkan memanfaatkan hasil penelitian ini untuk menyusun strategi peningkatan layanan berbasis data guna meningkatkan kepuasan dan pengalaman pengguna secara keseluruhan.

DAFTAR PUSTAKA

[1] Agneresa, A., Hananto, A. L., Hilabi, S. S., Hananto, A., & Tukino, T. (2022). Strategi Promosi Penerapan Data Mining Mahasiswa Baru Dengan Metode K-Means Clustering. *Dirgamaya: Jurnal Manajemen Dan Sistem Informasi*, *2*(2), 25–34. https://doi.org/10.35969/dirgamaya.v2i2.275

[2] Ding, K. and Jiang, P., 2017. RFID-based production data analysis in an IoT-enabled smart job-shop. *IEEE/CAA Journal of Automatica Sinica*. Dewi, N. L. P. P., Purnama, I. N., & Utami, N. W. (2022). Penerapan Data Mining Untuk Clustering Penilaian Kinerja Dosen Menggunakan Algoritma K-Means (Studi Kasus: STMIK Primakara). Jurnal Ilmiah Teknologi Informasi Asia, 16(2), 105. https://doi.org/10.32815/jitika.v16i2.761

[3] Ria, C. D., Gata, G., Hin, L. L., & Hamdani, A. U. (2024). Analisis Klasterisasi Data Mahasiswa Terhadap Kesehatan Mental Menggunakan Algoritma X-Means Clustering Analysis Of Student Data On Mental Health Using X-Means Algorithm. 3(September), 1571–1580.

[4] Sholeh, M., Suraya, S., & Andayati, D. (2024). Penerapan Data Mining pada Model Clustering Data Kuesioner Mahasiswa terhadap Kinerja

Dosen. Jurnal Eksplora Informatika, 13(2), 208–217. https://doi.org/10.30864/eksplora.v13i2.751

[5] Ginting, D. C. P., Sihombing, J. S. P., & ... (2023). Analisis Pemberian Insentif Tenaga Medis Menggunakan Algoritma K-Means Clustering. Jurnal Tekinkom …, 6, 213–219. https://doi.org/10.37600/tekinkom.v6i1.858

[6] Sulistyawati, A. A. D., & Sadikin, M. (2021). Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan. SISTEMASI: Jurnal Sistem …. https://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/view/1332

[7] Safitri, S. N., Haryono Setiadi, & Suryani, E. (2022). Educational Data Mining Using Cluster Analysis Methods and Decision Trees based on Log Mining. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 6(3), 448–456. https://doi.org/10.29207/resti.v6i3.3935

[8] Sajidah, S., Herdiana, R., & Solihudin, D. (2023). Segmentasi Pelanggan Salon Nuii Beauty Glow Menggunakan K-Means Clustering. JATI (Jurnal Mahasiswa Teknik Informatika), 7(1), 558–566. https://doi.org/10.36040/jati.v7i1.6333

[9] Sudrajat, A., Irma Purnamasari, A., & Ali, I. (2024). Penerapan K-Means Untuk Menganalisis Kepuasan Mahasiswa Terhadap Pembelajaran Daring. JATI (Jurnal Mahasiswa Teknik Informatika), 8(1), 1083–1089. https://doi.org/10.36040/jati.v8i1.8337

[10] Jannah, E. R., & Martanto. (2024). Klaterisasi Data Penduduk Berdasarkan Pekerjaan Menggunakan Metode K-Means Pada Wilayah Jawa Barat. JATI (Jurnal Mahasiswa Teknik Informatika), 8(2), 1709–1716. https://doi.org/10.36040/jati.v8i2.9055

[11] Nugraha, M. F., Martano, M., & Hayati, U. (2024). Clustering Data Indonesian Food Delivery Menggunakan Metode K-Means Pada Gofood Product List. JATI (Jurnal Mahasiswa Teknik …. https://ejournal.itn.ac.id/index.php/jati/article/view/9727

[12] Mar'atun Sholihah, O., Suarna, N., Dwilestari, G., & R, N. (2023). Implementasi Metode K-Means Clustering Untuk Menganalisa Penerima Bantuan Di Desa Palasah. Jurnal Informatika Dan Teknologi Informasi, 2(1), 154–160. https://doi.org/10.56854/jt.v2i1.123

[13] Patimah, E., Ermatita, E., & Chamidah, N. (2021). Analisis Cluster Kepuasan Pengguna Terhadap Layanan Shopee Menggunakan Algoritma K-Means. In Informatik : Jurnal Ilmu Komputer (Vol. 17, Issue 3). https://doi.org/10.52958/iftk.v17i3.3654

[14] Dayat, M. N., Suarna, N., & Wijaya, Y. A. (2023). Analisa Clustering untuk Mengelompokan Data Penayangan Film Bioskop Menggunakan Algoritma K-Means. INTERNAL (Information System Journal), 6(1), 68–78. https://doi.org/10.32627/internal.v6i1.686

[15] Tarigan, M. Y. B., Simanjuntak, M. A. S. B., Sitanggang, M. A. L. S., S. Prananta, E., & Sitanggang, D. (2023). Implementasi Data Mining Pengelompokan Tingkat Kepuasan Pelanggan Pada PT. Bisa Group Menggunakan Fuzzy C-Means. 6, 688–695. https://doi.org/10.37600/tekinkom.v6i2.937

[16] Wijaya, Y. A. (2021). Pengelompokan Data Menggunakan Davies-Bouldin Index. Journal of Data Science, 3(2), 115–130.

[17] Noor, M. (2024). Analisis Segmentasi Kepuasan Berdasarkan Usia. Journal of Customer Insights, 5(1), 45–60.

[18] Rahmawati, D. (n.d.). Peran Demografi dalam Analisis Data. Jurnal Teknologi Informasi, 7(3), 33–45.