

KATEGORISASI PELANGGAN BERDASARKAN TIPE DAN BIAYA CUCI KENDARAAN DENGAN MENERAPKAN ALGORITMA K-MEANS PADA SATRIA BIMA WASH

Fauhan Salsabila, Willy Prihartono

Komputerisasi Akuntansi, STMIK IKMI Cirebon

Jl. Perjuangan No. 10 B, Majasem, Cirebon Jawa Barat Indonesia

fauhansalsabila09@gmail.com

ABSTRAK

Satria Bima Wash adalah salah satu penyedia layanan cuci kendaraan yang memiliki berbagai data pelanggan, tetapi belum dimanfaatkan dengan maksimal. Studi ini bertujuan untuk menggunakan algoritma K-Means *Clustering* dalam mengelompokkan pelanggan berdasarkan tipe kendaraan, merek kendaraan, dan biaya layanan. Proses penelitian dimulai dengan pengumpulan data transaksi, dilanjutkan dengan *preprocessing* untuk membersihkan dan menormalkan data, dan diakhiri dengan evaluasi jumlah cluster yang optimal menggunakan *Davies-Bouldin Index* (DBI). Studi ini diharapkan dapat membantu pemilik bisnis dalam memahami perilaku konsumen, meningkatkan pelayanan, dan merancang strategi pemasaran yang lebih efisien. Hasil dari penelitian ini mengindikasikan bahwa pelanggan dapat dikategorikan menjadi tiga kelompok utama, dengan nilai DBI sebesar 0,005 yang menunjukkan mutu clustering yang sangat baik. Setiap cluster memiliki ciri khas tersendiri, seperti pilihan tipe kendaraan atau biaya layanan tertentu

Kata kunci : *Clustering, Algoritma K-Means, Pelanggan, Davies-Bouldin Index(DBI), Jasa Cuci Kendaraan, Strategi Pemasaran*

1. PENDAHULUAN

Pencucian kendaraan adalah salah satu jenis perawatan kendaraan bermotor dimana merupakan proses membersihkan kendaraan, baik itu mobil, motor, truk, atau kendaraan lainnya. Dalam mencuci kendaraan dapat dilakukan sendiri, atau menggunakan pelayanan pencucian kendaraan secara otomatis maupun secara manual dengan petugas yang membersihkan dari jasa cuci kendaraan tersebut.

Seiring dengan perkembangan dunia bisnis otomotif di negara Indonesia, kebutuhan transportasi di Indonesia terus meningkat, diikuti dengan bertambahnya jumlah penduduk tiap tahun. Dengan terus meningkatnya jumlah kendaraan berdampak positif pada usaha cuci mobil dan motor. Karena dengan adanya usaha cuci mobil atau motor bisa membantu, pelanggan yang sibuk dengan rutinitas, bisa mencuci kendaraan ke salah satu tempat cuci motor. Satria Bima Wash merupakan perusahaan yang bergerak dibidang jasa cuci kendaraan, perusahaan ini merupakan cabang kedua. Melihat pada cabang pertamanya yang sangat diminati oleh para pelanggan, yang membuat pemilik membuka cabang kedua. Perawatan kendaraan memang susah-susah gampang, jika bukan dengan ahlinya mungkin saja kendaraan tidak semenarik pada saat awal membelinya. Jasa adalah usaha yang menjual layanan kepada konsumen yang membutuhkan. Jasa adalah penyediaan produk yang tidak terlihat dan tidak terwujud dan tidak menyebabkan konsumen mendapat hak milik atas produk tersebut. Namun dalam jasa, konsumen hanya mendapatkan pelayanan yang sesuai atau tidak sesuai dengan harapannya setelah mereka melakukan transaksi. Hal tersebut yang membuat jasa sangat

berbeda dengan produk. Oleh karenanya cara dan metode untuk melakukan pemasarannya juga berbeda. [1]

Analisis lingkungan usaha adalah proses untuk memahami lingkungan bisnis perusahaan, termasuk kekuatan dan kelemahan perusahaan, serta peluang dan ancaman yang ada di lingkungan tersebut. Analisis ini bertujuan untuk membantu perusahaan dalam mengidentifikasi peluang dan ancaman, mengembangkan strategi bisnis, dan meningkatkan efektivitas manajemen[2]. Strategi pemasaran adalah membuat keputusan tentang biaya pemasaran dari anggaran pemasaran, dengan mempertimbangkan kondisi lingkungan dan persaingan yang diantisipasi. [3]

Penelitian ini menggali aspek-aspek tersebut dengan menyelidiki bagaimana normalisasi dataset yang dikelompokkan memengaruhi perhitungan jarak Euclidean dalam algoritma K-means. Selain itu, penelitian ini menjelajahi pengaruh berbagai jumlah kluster, yang ditentukan melalui metode elbow, terhadap evaluasi Indeks Davies Bouldin.[4]

Tujuan penelitian ini adalah untuk menerapkan Algoritma K-Means dalam menganalisis dan mengelompokkan data transaksi dan pola perilaku pelanggan. Dengan segmentasi ini diharapkan pemilik usaha dapat memahami pola perilaku pelanggan dan mengoptimalkan pengelolaan sumber daya serta pelayanan

2. TINJAUAN PUSTAKA

2.1. K-Means Clustering

K-Means *Clustering* merupakan suatu metode untuk mencari dan mengelompokkan data yang

memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. *Clustering* merupakan salah satu metode data mining yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target output.[5]

Clustering adalah peroses pengelompokan objek ke dalam class dengan mengetahui kesamaan objek. Sebuah *cluster* adalah kumpulan data yang mempunyai persamaan satu dengan yang lain, dalam *cluster* yang mempunyai kesamaan dan memiliki perbedaan dengan *cluster* yang lain. *Clustering* merupakan pembelajaran yang tidak dapat ditebak karena tidak dibutuhkan target *output* atau yang sering disebut *unsupervised learning*. [6]

2.2. Transaksi

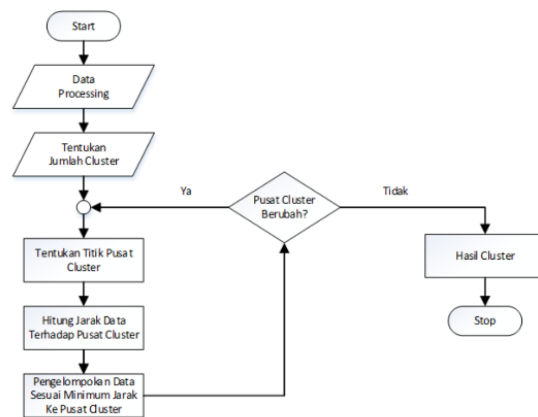
Transaksi adalah suatu kegiatan yang dilakukan seseorang yang menimbulkan perubahan terhadap harta atau keuangan yang dimiliki baik itu bertambah ataupun berkurang. Misalnya menjual harta, membeli barang, membayar hutang, serta membayar berbagai mava biaya untuk memenuhi sebuah kebututuhan hidup.[7]

2.3. Pelanggan

Pelanggan adalah orang yang secara tetap menggunakan produk atau jasa kita serta ikut membantu mengarahkan serta menuntut suatu perusahaan atau lembaga untuk berbuat sesuai dengan standar-standar kualitas tertentu menurut sebahagian keinginan mereka. Tuntutan dan keinginan mereka tersebut ikut mempengaruhi penampilan atau kebijakan suatu lembaga. Lembaga tersebut secara otomatis kemudian bergantung pada keinginan dan kebutuhan para pelanggan, sehingga hampir semua keinginan pelanggan adalah menjadi bagian prioritas yang harus dipenuhi.[8]

2.4. Data Mining

Data mining adalah suatu proses ekstraksi atau penggalian data dan informasi yang besar, yang belum diketahui sebelumnya, namun dapat dipahamidan berguna dari *database* yang besar serta digunakan untuk membuat suatu keputusanbisnis yang sangat penting. *Data mining* menggambarkan sebuah pengumpulan teknik-teknik dengan tujuan untuk menemukan pola-pola yang tidak diketahui pada data yang telah dikumpulkan. *Data mining* adalah bagian dari proses KDD (*Knowledge Discovery in Databases*) yang terdiri dari beberapa tahapan seperti pemilihan data, pra pengolahan, transformasi, *data mining*, dan evaluasi hasil. KDD secara umum juga dikenal sebagai pangkalan data.[9]



Gambar 1. Arus Algoritma K-Means Clustering

3. METODE PENELITIAN

3.1. Sumber Data

Sumber data ini diperoleh dari operasional harian yang dilakukan di cabang kedua Satria Bima Wash. Lokasi cabang ini berada di Jalan Bima Raya, Sunyaragi, Kec. Kesambi, Kota Cirebon yang merupakan tempat utama dalam pengumpulan data penelitian ini. Perolehan data ini di dapatkan dengan wawancara dengan salah satu pemilik Satria Bima Wash. Data dikumpulkan selama periode Maret hingga Mei 2024, dengan rentang waktu yang cukup untuk mengembangkan pola layanan pencucian kendaraan.

3.2. Teknik Pengumpulan Data

Teknik pengumpulan data adalah metode untuk mendapatkan informasi yang relevan tentang penelitian ini. Adapun teknik pengumpulan data yang digunakan adalah:

a. Wawancara

Penelitian ini melakukan wawancara langsung dengan pemilik cabang kedua Satria Bima Wash yaitu Pak Jafar, serta beberapa karyawan. Tujuannya adalah mengetahui layanan apa saja yang diminati, kendala yang sering dihadapi, dan proses operasionalnya. Data dari wawancara ini akan menjadi dasar untuk analisis penelitian.

b. Observasi

Observasi dilakukan dalam penelitian untuk mengamati proses operasional di Satria Bima Wash.

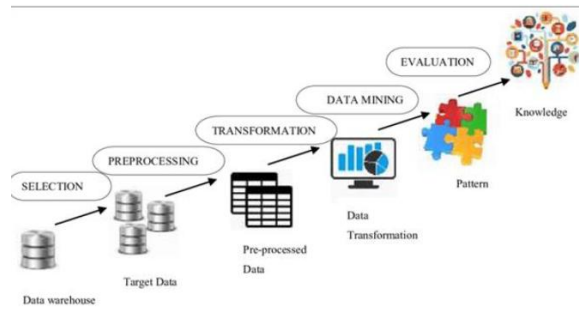
c. Dokumentasi

Dokumentasi dilaksanakan dengan maksud melihat dan mencatat laporan atau arsip yang mendukung data untuk penelitian.

3.3. Tahap Perancangan

Pada proses tahapan perancangan data ini berdasarkan *Knowledge Discovery in Database* (KDD) adalah proses analisa terstruktur untuk memperoleh informasi yang benar, baru, bermanfaat dan menemukan pola dari data yang besar dan komplek.[10]

Adapun metode yang digunakan dalam pelaksanaan penelitian ini, seperti gambar yang ada dibawah ini.



Gambar 2. Tahapan Penelitian Menggunakan KDD

Langkah-langkah Metodologi *Knowledge Discovery in Database* (KDD) dijelaskan sebagai berikut:

a. *Data Selection*

Tahap pertama adalah seleksi data. Pada fase ini, data yang relevan dipilih dari sumber data primer agar data yang akan diolah sesuai dengan tujuan penelitian yang ini dicapai. Proses ini menyaring data yang dianggap penting untuk analisis lebih lanjut.

b. *Preprocessing*

Setelah data dipilih, proses selanjutnya adalah preprocessing data atau pra-pemrosesan. Data yang dipilih sering kali memiliki ketidaksempurnaan seperti nilai yang hilang, duplikat, dan ketidakkonsistenan. Pada fase ini, data dibersihkan dan meningkatkan kualitasnya. Untuk memastikan data dapat digunakan pada fase selanjutnya.

c. *Transformation*

Tahap Transformation atau transformasi data adalah tahapan dimana data yang telah dibersihkan diubah menjadi format yang lebih sesuai untuk analisis. Transformasi ini mungkin melibatkan pengumpulan data, pengkodean ulang atribut, atau pembuatan atribut baru.

d. *Data Mining*

Setelah data diproses, fase inti proses KDD dimulai yaitu Data mining. Disini Algoritma tertentu digunakan untuk menemukan pola dan model dalam data. Algoritma yang digunakan bergantung pada tujuan analisis. Selama fase ini, analisis terperinci dilakukan untuk menemukan pola yang tersembunyi.

e. *Evaluation*

Setelah pola atau model ditemukan, langkah selanjutnya adalah evaluasi. Pada tahap ini, hasil analisis dievaluasi untuk menentukan apakah pola atau model yang dihasilkan sesuai dengan tujuan penelitian yang ingin dicapai.

f. *Knowledge*

Fase terakhir adalah fasa *Knowledge*. Disini, pola-pola yang dievaluasi diubah menjadi pengetahuan yang dapat digunakan untuk mengambil keputusan atau memberikan wawasan baru. Pengetahuan ini

dapat membantu pemilik usaha untuk menentukan strategi bisnis atau mengidentifikasi peluang baru.

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

Pemilihan data, yang merupakan tahap dimana memilih atribut yang akan digunakan untuk dianalisis terhadap data yang relevan.

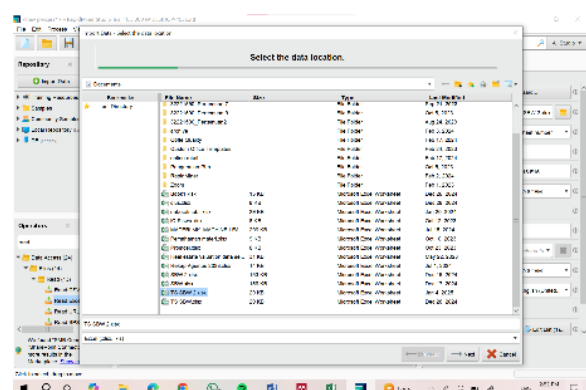
Tabel 1. Dataset

Tanggal	Jenis Kendaraan	Merk Kendaraan	Biaya
01-04-2024	Motor	Yamaha NMAX	15.000
01-04-2024	Motor	Honda Beat	15.000
02-04-2024	Mobil	Toyota Vios	35.000
02-04-2024	Mobil	Misubishi Xpander	35.000
03-04-2024	Motor	Honda Vario	10.000
03-04-2024	Mobil	Toyota Hiace	40.000
04-04-2024	Mobil	Mitsubishi Pick up	35.000

4.2. Preprocessing

Proses Preprocessing dilanjutkan dengan memasukkan dataset yang terdiri dari 673 transaksi dari bulan Maret sampai bulan Mei 2024. Data tersebut tersimpan dalam Microsoft Excel dan akan dikelola menggunakan perangkat lunak rapidminer.

Gambar 2. Dataset Excel

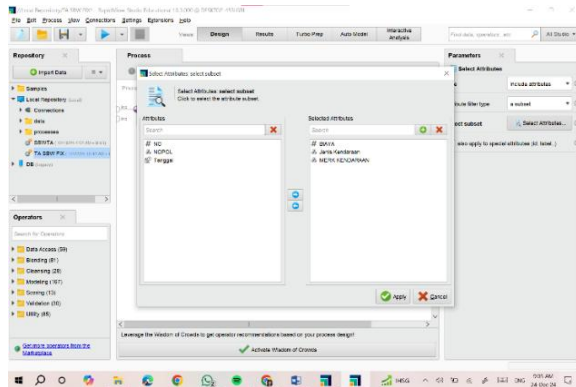


Gambar 3. Proses Memasukkan Data

4.3. Transformation

Tahap transformasi merupakan langkah berikutnya setelah *preprocessing*. Pada tahap ini, data diubah sesuai kebutuhan analisis atau pemodelan yang akan diterapkan. Dalam Algoritma K-Means

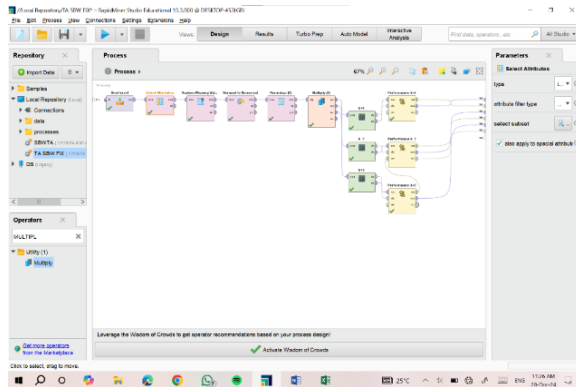
transformasi adalah dimana data yang telah dibersihkan diubah menjadi format yang lebih sesuai untuk analisis.



Gambar 4. Tahapan transformasi

4.4. Data Mining

Teknik *Data mining* menyimpan data untuk menghasilkan informasi yang berguna. Pada tahap penelitian ini, *data mining* dengan menggunakan metode Algoritma K-Means Clustering untuk mengetahui pengelompokan pelanggannya. Untuk dapat melakukan proses data yang sesuai dengan teknik *data mining* maka digunakan operator-operator yang telah tersedia pada tools RapidMiner, masing-masing operator tersebut memiliki fungsi yang berbeda-beda.



Gambar 5. Perancangan Pada RapidMiner

4.5. Evaluation

Untuk mengetahui jumlah kelompok yang optimal dengan menggunakan metode K-Means maka menggunakan operator *cluster distance performance* yang berfungsi untuk mengetahui nilai *Davies-Bouldin Index* (DBI).

Tabel 2. Parameter Cluster Distance Performance

K	DBI
1	0.007
2	0.006
3	0.005

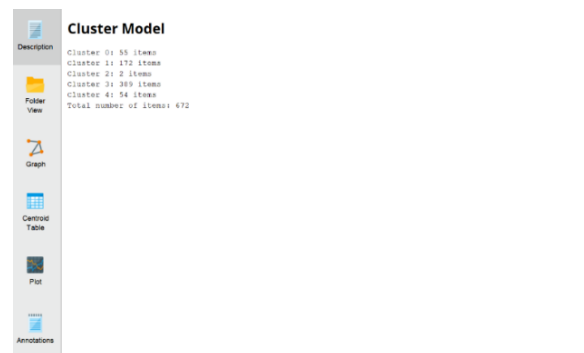
Berdasarkan hasil yang diperoleh dalam mencari *cluster* yang optimal menggunakan *cluster distance*

performance, bahwa pengelompokan terbaik adalah *cluster 3* dengan nilai *Davies-Bouldin Index* (DBI) sebesar 0.005 nilai tersebut paling optimal karena mendekati angka 0.

Tabel 3. Jarak Avg Within Centroid Distance

0	1	2	3	4
0.002	0.005	0.058	0.006	0.000

Dari table 3 menunjukkan antara jarak *avg within centroid distance* dengan *centroid distance per cluster*. Dimana hasil terendah menunjukkan bahwa *cluster* tersebut adalah *cluster 0*.



Gambar 6. Hasil Cluster Model

Pada gambar 4.18 menunjukkan pembagian data pelanggan ke dalam 5 *cluster*. *Cluster 0* terdapat 55 pelanggan, *cluster 1* terdapat 172 pelanggan, *cluster 2* terdapat 2 pelanggan, *cluster 3* terdapat 389 pelanggan, dan *cluster 4* terdapat 54 pelanggan.

Tabel 4. Hasil Pengelompokan

Cluster	items
0	55
1	172
2	2
3	389
4	54

4.6. Knowledge

Penelitian ini menghasilkan pengelompokan pelanggan ke dalam 5 *cluster* menggunakan Algoritma K-Means Clustering, dengan nilai *Davies-Bouldin Index* (DBI) sebagai indikator *cluster*. Semakin rendah nilai DBI semakin baik kualitas *cluster* yang terbentuk. Didapatkan percobaan dengan

4.7. K-2

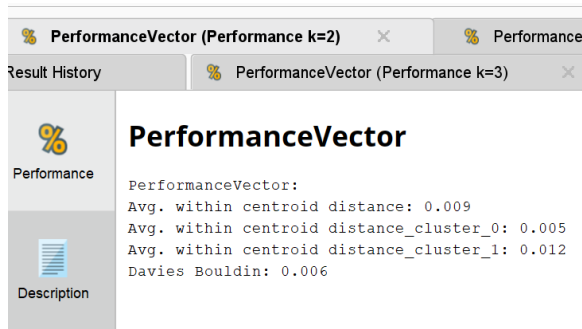
Criterion
Avg. within centroid dis...
Avg. within centroid dis...
Avg. within centroid dis...
Davies Bouldin

Davies Bouldin

Davies Bouldin: 0.006

Gambar 7. Hasil DBI K-2

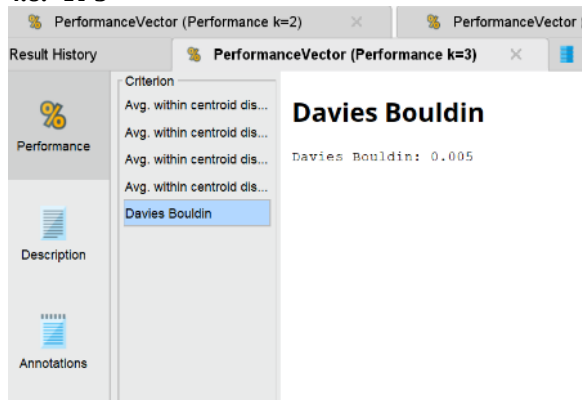
Hasil perhitungan *Davies-Bouldin Index* pada gambar 7 menunjukkan nilai sebesar 0.006 yang mengidentifikasi kualitas *clustering* yang sangat baik.



Gambar 8. Hasil Jarak Centroid

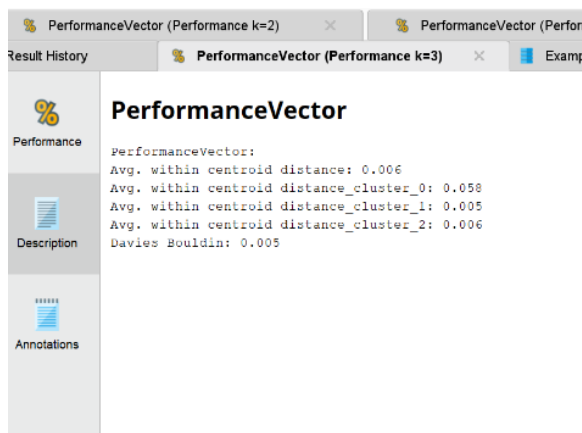
Hasil evaluasi pada gambar 8 menunjukkan performa clustering dengan menggunakan $k=2$. Nilai rata-rata jarak antar *cluster* terhadap centroid nya (*Avg. within centroid*) adalah 0.009 yang berarti data dalam cluster memiliki tingkat kedekatan yang cukup tinggi dengan centroid.

4.8. K-3



Gambar 9. Hasil DBI K-3

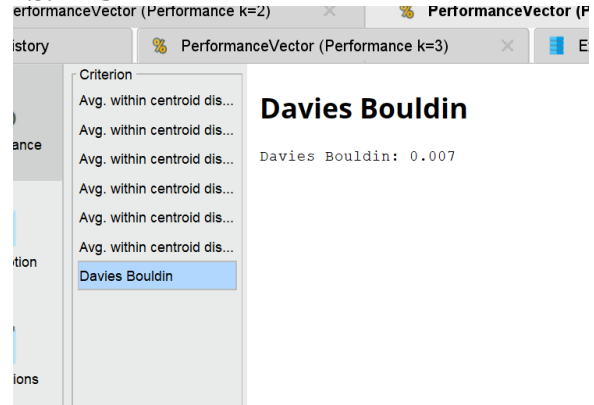
Hasil perhitungan *Davies-Bouldin Index* pada gambar 9 menunjukkan nilai sebesar 0.005 yang mengidentifikasi kualitas *clustering* yang sangat baik.



Gambar 10. Hasil Jarak Centroid

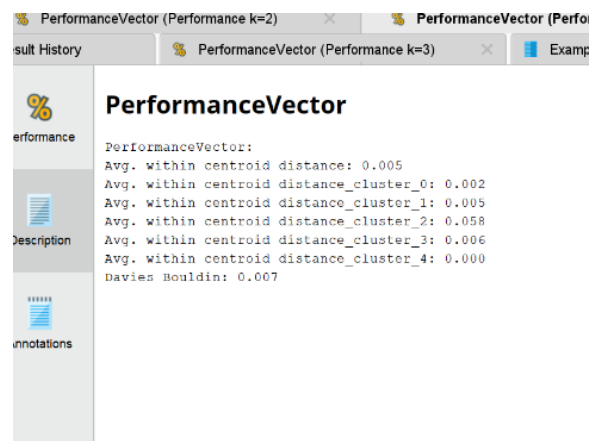
Hasil evaluasi pada gambar 10 menunjukkan performa clustering dengan menggunakan $k=3$. Nilai rata-rata jarak antar cluster terhadap centroid nya (*Avg. within centroid*) adalah 0.005 yang berarti data dalam cluster memiliki tingkat kedekatan yang cukup tinggi dengan centroid.

4.9. K-5



Gambar 11. Hasil DBI K-5

Hasil perhitungan *Davies-Bouldin Index* pada gambar 11 menunjukkan nilai sebesar 0.007 yang mengidentifikasi kualitas *clustering* yang cukup baik.



Gambar 12. Hasil Jarak Centroid

Hasil pada gambar 12 menunjukkan data di data *cluster* yang sangat dekat dengan *centroid*. Berdasarkan hasil clustering disajikan tabel per-clusternya sebagai berikut:

4.10. Cluster 0 terdapat 2 pelanggan

Tabel 5. Hasil Cluster 0

id	Jenis Kendaraan	Merek Kendaraan	Biaya	Kategori
341	Mobil	Mitsubishi Pajero	Diatas 40.000	Layanan jarang dipilih
342	Mobil	Toyota Crown	Diatas 40.000	Layanan jarang dipilih

4.11. Cluster 1 terdapat 281 pelanggan

Tabel 6. Hasil Cluster 1

id	Jenis Kendaraan	Merek Kendaraan	Biaya	Kategori
1	Motor	Yamaha Nmax	standar	Sering dipilih
2	Motor	Honda Beat	standar	Sering dipilih
4	Motor	Vespa	Standar	Sering dipilih
671	Motor	Vario	standar	Sering dipilih

4.12. Cluster 2 terdapat 389 pelanggan

Tabel 7. Hasil Cluster 2

id	Jenis Kendaraan	Merek Kendaraan	Biaya	Kategori
3	Mobil	Mitsubishi Xpander	standar	Sering dipilih
5	Motor	Honda Beat	standar	Sering dipilih
9	Mobil	Daihatsu Xenia	standar	Sering dipilih
672	Motor	Toyota Etios	standar	Sering dipilih

5. KESIMPULAN DAN SARAN

Hasil penelitian ini menunjukkan bahwa penerapan Algoritma K-Means *Clustering* berhasil mengelompokkan pelanggan jasa cuci kendaraan ke dalam tiga *cluster* berdasarkan jenis kendaraan, merek kendaraan, dan biaya layanan, dengan masing-masing *cluster* memiliki karakteristik unik. Evaluasi menggunakan *Davies-Bouldin Index* (DBI) menunjukkan kualitas clustering yang sangat baik, dengan nilai DBI sebesar 0,005, yang mengindikasikan jarak antar *cluster* yang signifikan dan homogenitas data dalam *cluster*. Pengelompokan ini memberikan wawasan bagi pemilik usaha dalam memahami pola pelanggan, sehingga dapat digunakan untuk mengatur strategi promosi dan harga yang lebih kompetitif. Sebagai saran, penelitian ini diharapkan dapat membantu pemilik usaha dalam mengembangkan strategi pemasaran yang lebih efektif dengan

DAFTAR PUSTAKA

- [1] O. Hapsara et al., *Manajemen Pemasaran Jasa : Membangun Loyalitas Pelanggan*. 2023.
- [2] A. C. Adjie, "Analisis Lingkungan Dan Kelayakan Usaha Salon Mobil Deluxe Auto Detailer Semarang," *Sanskara Ekon. dan Kewirausahaan*, vol. 2, no. 02, pp. 111–117, 2024, doi: 10.58812/sek.v2i02.317.
- [3] D. N. S. Rambe and N. Aslami, "Analisis Strategi Pemasaran Dalam Pasar Global," *El-Mujtama J. Pengabd. Masy.*, vol. 1, no. 2, pp. 213–223, 2022, doi: 10.47467/elmujtama.v1i2.853.
- [4] I. T. Umagapi, B. Umaternate, S. Komputer, P. Pasca Sarjana Universitas Handayani, B. Kepegawaian Daerah Kabupaten Pulau Morotai, and B. Riset dan Inovasi, "Uji Kinerja K-Means Clustering Menggunakan Davies-Bouldin Index Pada Pengelompokan Data Prestasi Siswa," *Semin. Nas. Sist. Inf. dan Teknol.*, vol. 7, no. 1, pp. 303–308, 2023.
- [5] R. Supardi and I. Kanedi, "Implementasi Metode Algoritma K-Means Clustering pada Toko Eidelweis," *J. Teknol. Inf.*, vol. 4, no. 2, pp. 270–277, 2020, doi: 10.36294/jurti.v4i2.1444.
- [6] Z. Muttaqin, "Implementasi Unsupervised Learning Pada Nilai Jasmani Kesamaptaan Sekolah Polisi Negara Dengan Metode Clustering Analysis," *PROSISKO J. Pengemb. Ris. dan Obs. Sist. Komput.*, vol. 10, no. 1, pp. 18–23, 2023, doi: 10.30656/prosisko.v10i1.6269.
- [7] E. I. Syaripudin, D. Konkon Furkony, M. Maulin, and H. Bisri, "Prinsip-Prinsip Dan Kaidah Transaksi Dalam Ekonomi Syari'Ah," *J. Huk. Ekon. Syariah*, vol. 1, no. 2, pp. 284–294, 2023, doi: 10.37968/jhesy.v1i2.359.
- [8] Irwasyah and M. Suradji, "Pelanggan Dan Karakteristiknya," *J. Stud. Pendidik. Islam*, vol. 4, no. 2, pp. 170–188, 2021.
- [9] C. Zai, "Implementasi Data Mining Sebagai Pengolahan Data," *J. Portal Data*, vol. 2, no. 3, pp. 1–12, 2022, [Online]. Available: <http://portaldata.org/index.php/portaldata/article/view/107>
- [10] Y. Pratiwi and M. Mulyawan, "Implementasi Algoritma K-Means untuk Menentukan Angka Harapan Hidup berdasarkan Tingkat Provinsi," *Blend Sains J. Tek.*, vol. 1, no. 4, pp. 284–294, 2023, doi: 10.56211/blendsains.v1i4.233.