

Implementasi K-Means Clustering Untuk Mengelompokkan Data Sparepart Alat Berat

Andira Rahmawati¹, Muhammad Fadlan², Anto³

^{1,2,3} Sistem Informasi, STMIK PPKIA Tarakanita Rahmawati, Tarakan, Kalimantan Utara
Email: ¹andiraaa2@gmail.com, ²fadlan @ppkia.ac.id, ³anto@ppkia.ac.id

Abstrak

Data mining merupakan proses penting dalam mengekstraksi informasi yang bernilai dari data yang ada, yang kemudian dapat digunakan oleh perusahaan untuk pengambilan keputusan yang cepat dan akurat. Salah satu metode yang sering digunakan dalam data mining adalah metode K-Means Clustering. Dalam penelitian ini, penulis menerapkan metode K-Means Clustering di bidang retail untuk mengatasi permasalahan yang dihadapi oleh PT. Patria Jaya Mandiri. Penulis merancang sebuah aplikasi yang dapat mengelompokkan sparepart alat berat berdasarkan data penjualan, dengan tujuan membantu perusahaan dalam mengetahui sparepart mana saja yang paling diminati oleh konsumen. Dengan adanya pengelompokan ini, diharapkan perusahaan dapat lebih mudah menentukan stok sparepart yang optimal, yang pada akhirnya akan berdampak positif pada pendapatan perusahaan. Hasil dari penelitian ini menunjukkan bahwa sparepart alat berat dapat dikelompokkan menjadi tiga kategori: Paling Laris, Laris, dan Tidak Laris. Klaster 1 (Paling Laris) terdiri dari 3 data, klaster 2 (Laris) terdiri dari 39 data, dan klaster 3 (Tidak Laris) terdiri dari 8 data. Hasil clustering ini dapat menjadi panduan bagi PT. Patria Jaya Mandiri dalam menentukan persediaan sparepart yang tepat di masa mendatang.

Kata Kunci: centroid, data mining, k-means clustering, pengelompokan, sparepart.

K-Means Clustering Implementation For Grouping Heavy Equipment Spareparts Data

Abstract

Data mining is a crucial process for extracting valuable information from existing data, which can then be used by companies for quick and accurate decision-making. One of the commonly used methods in data mining is the K-Means Clustering method. In this study, the author applied K-Means Clustering in the retail sector to address the challenges faced by PT. Patria Jaya Mandiri. The author designed an application that can cluster heavy equipment spare parts based on sales data, with the aim of helping the company identify which spare parts are most favored by consumers. This clustering is expected to simplify the process of determining optimal spare part stock, ultimately positively impacting the company's revenue. The results of this study indicate that heavy equipment spare parts can be categorized into three groups: Most Popular, Popular, and Least Popular. Cluster 1 (Most Popular) consists of 3 data points, Cluster 2 (Popular) consists of 39 data points, and Cluster 3 (Least Popular) consists of 8 data points. This clustering result can serve as a guide for PT. Patria Jaya Mandiri in determining the optimal spare part inventory in the future.

Keywords: centroid, data mining, grouping, k-means clustering, sparepart.

I. PENDAHULUAN

Kebutuhan manusia akan informasi terus meningkat seiring dengan pesatnya perkembangan ilmu pengetahuan dan teknologi. Teknologi memungkinkan manusia untuk menciptakan berbagai alat yang mendukung kehidupan, salah satunya adalah komputer. Komputer sangat membantu dalam menyelesaikan berbagai pekerjaan, terutama dalam pengolahan data, sehingga dapat menghasilkan informasi

yang dibutuhkan oleh organisasi atau perusahaan. Dengan demikian, organisasi dan perusahaan harus mampu menyesuaikan diri dengan perkembangan teknologi yang ada saat ini [1].

Salah satu masalah penting yang dihadapi perusahaan adalah pengendalian persediaan stok barang, khususnya terkait kekurangan stok. Kekurangan stok dapat mengecewakan pelanggan, yang pada gilirannya dapat menurunkan pendapatan perusahaan. Salah satu solusi untuk

masalah ini adalah dengan menerapkan pendekatan data mining untuk menganalisis data penjualan yang ada. Dengan memanfaatkan data mining, perusahaan dapat menggali informasi penting dari data penjualan historis, yang dapat digunakan untuk memprediksi pola permintaan di masa depan dan mengoptimalkan manajemen persediaan.

PT. Patria Jaya Mandiri, sebuah perusahaan yang bergerak di bidang jual beli suku cadang alat berat di Kota Tarakan, menghadapi masalah serupa. Perusahaan ini sering kali kesulitan untuk mengetahui suku cadang mana yang paling diminati konsumen, sehingga kadang terdapat suku cadang yang tidak laku terjual. Kondisi ini dapat menyebabkan kerugian bagi perusahaan, terutama karena suku cadang yang tidak laku berpotensi menjadi usang atau rusak, sehingga nilainya menurun drastis.

Untuk mengatasi masalah ini, PT. Patria Jaya Mandiri perlu mengidentifikasi suku cadang yang paling diminati konsumen agar dapat memperbanyak stok suku cadang tersebut, sekaligus mengurangi stok suku cadang yang kurang diminati. Identifikasi ini dapat dilakukan dengan menggunakan teknik data mining seperti K-Means Clustering, yang mampu mengelompokkan data penjualan ke dalam beberapa kategori berdasarkan pola pembelian konsumen. Dengan demikian, perusahaan dapat lebih efisien dalam mengelola persediaan, mengurangi risiko kerugian akibat suku cadang yang tidak terjual, serta memastikan bahwa suku cadang yang paling dibutuhkan selalu tersedia. Langkah ini tidak hanya dapat meningkatkan kepuasan pelanggan tetapi juga menjaga stabilitas pendapatan perusahaan.

Data mining merupakan proses penggalian informasi penting dari sejumlah data. Data mining melibatkan serangkaian proses untuk menggali nilai tambah berupa informasi yang tidak dapat diketahui secara manual dari basis data. Informasi yang dihasilkan diperoleh dengan mengekstraksi dan mengenali pola penting dari data yang ada di dalam basis data. Data mining dapat membantu perusahaan dalam pengambilan keputusan dengan cepat dan tepat. Salah satu metode dalam data mining adalah K-Means Clustering [2].

K-Means Clustering adalah teknik data mining yang menggunakan pendekatan unsupervised learning. Metode ini bertujuan untuk mengelompokkan data yang memiliki karakteristik serupa ke dalam satu kelompok atau cluster yang sama, sementara data dengan karakteristik berbeda dikelompokkan ke dalam cluster lain [3]. Dalam praktiknya, metode ini cukup mudah dan cepat diimplementasikan, serta memiliki kemampuan yang besar dalam mengolah data dalam jumlah besar dengan waktu yang lebih efisien. Beberapa penelitian menggunakan K-Means telah digunakan sebelumnya, diantaranya penelitian oleh Hardiani (2022) menggunakan K-Means untuk klasterisasi kasus Covid-19 di Indonesia [4], penelitian oleh Marpaung (2023) menggunakan K-Means untuk klasterisasi data penduduk pada sebuah kabupaten di Indonesia [5], selain itu K-Means juga digunakan untuk klasterisasi provinsi yang ada di Indonesia berdasarkan potensi desa yang ada dalam provinsi tersebut [6].

Dalam penelitian ini penulis mencoba menerapkan metode K-Means Clustering di bidang retail untuk memecahkan masalah yang dialami oleh PT. Patria Jaya Mandiri dengan merancang sebuah aplikasi yang dapat melakukan pengelompokan sparepart alat berat dengan menggunakan K-Means Clustering. Data penjualan sparepart tersebut perlu dikelompokkan agar dapat membantu perusahaan dalam mengetahui sparepart apa saja yang diminati oleh konsumen, sehingga dengan adanya aplikasi ini diharapkan juga dapat memudahkan perusahaan dalam menentukan stok persediaan sparepart yang berpengaruh pada pendapatan perusahaan.

II. METODOLOGI PENELITIAN

Pada sub bagian metodologi penelitian, dijelaskan secara rinci mengenai alur atau tahapan penelitian yang dilakukan untuk mencapai tujuan yang telah ditetapkan. Metodologi ini mencakup langkah-langkah sistematis yang diterapkan mulai dari tahap awal hingga akhir penelitian. Setiap tahapan dirancang untuk memastikan bahwa penelitian berjalan secara terstruktur, efisien, dan menghasilkan temuan yang valid serta dapat diandalkan. Melalui metodologi ini, diharapkan penelitian dapat memberikan solusi yang tepat dan aplikatif terhadap permasalahan yang dihadapi. Tahapan atau alur penelitian dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

A. Observasi

Observasi merupakan langkah awal dalam penelitian ini di mana peneliti melakukan pengamatan langsung terhadap PT. Patria Jaya Mandiri, yang berfungsi sebagai objek penelitian. Melalui observasi ini, peneliti mendapatkan pemahaman yang lebih mendalam mengenai kondisi dan permasalahan yang dihadapi oleh perusahaan, khususnya terkait pengendalian persediaan stok barang.

B. Pengumpulan Data

Proses pengumpulan data dalam penelitian ini mencakup beberapa langkah penting. Pertama, peneliti mengambil data

langsung dari PT. Patria Jaya Mandiri yang relevan dengan topik penelitian. Data ini mencakup informasi penjualan, stok barang, dan pola permintaan konsumen. Selain itu, peneliti juga mengumpulkan data referensi dari berbagai sumber literatur seperti buku, jurnal, dan website yang berkaitan dengan metode K-Means Clustering. Data referensi ini digunakan untuk memperkaya landasan teori serta memperkuat analisis dalam penelitian.

C. Analisa Data dan Perancangan Sistem

Pada tahap ini, peneliti melakukan analisis terhadap data yang telah dikumpulkan dan merancang sistem yang sesuai dengan kebutuhan perusahaan. Analisis ini melibatkan pemahaman mendalam terhadap teori-teori yang digunakan dalam penelitian dan penerapan metode yang sesuai. Berikut adalah beberapa poin penting dalam tahap ini:

1. Data Mining

Data mining adalah proses mengekstraksi informasi berharga dari sejumlah besar data yang sebelumnya tidak diketahui namun dapat dipahami dan bermanfaat. Informasi ini kemudian digunakan untuk mendukung keputusan bisnis yang penting. Data mining tidak hanya mengungkap pola-pola tersembunyi dalam data, tetapi juga membantu organisasi dalam memahami tren dan perilaku yang relevan, sehingga memungkinkan pengambilan keputusan yang lebih tepat dan strategis.

2. K-Means

K-Means adalah salah satu metode clustering yang paling umum digunakan dalam data mining. Metode ini bekerja dengan cara mengelompokkan data ke dalam cluster berdasarkan kemiripan karakteristiknya. Clustering adalah proses membagi satu set objek data menjadi himpunan bagian yang disebut cluster, di mana objek-objek dalam satu cluster memiliki kemiripan satu sama lain lebih besar dibandingkan dengan objek-objek di cluster lain. Algoritma K-Means bekerja dengan menemukan titik pusat dari setiap cluster, yang disebut centroid, dan mengelompokkan data berdasarkan kedekatannya dengan centroid tersebut [7].

Tujuan utama dari K-Means adalah untuk memaksimalkan kemiripan data dalam satu cluster dan meminimalkan kemiripan data antar cluster. Fungsi jarak, seperti Euclidean distance, digunakan untuk mengukur kemiripan antara data dan centroid, di mana jarak yang lebih pendek menunjukkan kemiripan yang lebih besar. Algoritma K-Means sangat efektif dalam mengelompokkan data karena cepat, mudah diimplementasikan, dan dapat beradaptasi dengan berbagai jenis data. Algoritma ini menemukan kelompok-kelompok non-overlapping dalam data, yang membuatnya sangat berguna dalam berbagai aplikasi, termasuk pengelompokan dokumen atau objek dalam dataset besar [2], [3].

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (1)$$

Keterangan :

d = Distance

x = Data testing

y = Data training

xi = Data testing ke -i

yi = Data training ke -i

Proses klasterisasi dapat dilakukan dengan cara mengidentifikasi data yang akan diklaster, setelah melakukan identifikasi data, selanjutnya dapat menghitung nilai jarak antara data ke x dan data ke y. Dimana telah diketahui bahwa xi merupakan data testing ke-i sedangkan yi adalah data training ke-i, selanjutnya mencari nilai titik centroid dari klaster ke k sesuai rumus berikut ini.

$$\mu_k = \frac{1}{N_k} \sum_{i=R1}^{N_k} x_i \quad (2)$$

Keterangan :

μ_k = Nilai titik centroid

N_k = Banyak data cluster

xi = Nilai data ke - i

Berdasarkan rumus diatas dapat disimpulkan bahwa μ_k adalah nilai titik centroid dari cluster ke K. Setelah itu, simbol N_k berfungsi untuk menunjukkan banyaknya data pada cluster ke K. Selanjutnya simbol xi digunakan untuk menunjukkan nilai data ke i pada cluster ke K. Dataset yang telah final dapat dilakukan pengolahan data menggunakan Algoritma K-Means. Jumlah nilai K dapat ditentukan dengan membentuk jumlah cluster, lalu mencari titik centroid, dan pengelompokkan data berdasarkan jumlah cluster yang sudah ditentukan. Jika titik centroid berubah maka dilakukan lagi pemilihan titik centroid hingga diketahui tidak ada perubahan pada titik centroid.

D. Uji Coba

Tahap ini, pengujian aplikasi sedang dilakukan dengan menggunakan metode K-Means untuk memastikan masalah yang ada dapat teratasi dengan baik.

III. HASIL DAN PEMBAHASAN

A. Analisa Data

Data yang didapatkan oleh PT. Patria Jaya Mandiri sebanyak 50 data. Tujuan dari analisis ini adalah untuk membagi data tersebut ke dalam tiga cluster yang menggambarkan kategori penjualan: cluster 1 (C1) sebagai penjualan paling laris, cluster 2 (C2) sebagai penjualan laris, dan cluster 3 (C3) sebagai penjualan tidak laris. Untuk mencapai tujuan ini, digunakan metode K-Means yang memungkinkan pengelompokan data berdasarkan karakteristik penjualan yang berbeda.

B. Implementasi K-Means

Tahap awal dalam penerapan metode K-Means adalah menentukan jumlah cluster yang diinginkan. Dalam penelitian ini, diputuskan untuk membentuk tiga cluster. Langkah selanjutnya adalah menetapkan titik pusat awal (centroid) untuk setiap cluster. Centroid ini berfungsi sebagai titik acuan untuk mengelompokkan data. Pada awalnya, centroid dipilih secara acak dari data yang ada. Pemilihan centroid awal ini penting karena akan mempengaruhi hasil akhir clustering, dan proses iterasi dilakukan untuk memperbarui centroid hingga

stabilitas tercapai, yaitu saat tidak ada perubahan signifikan pada posisi centroid. Dengan demikian, setiap data penjualan dapat dikategorikan secara akurat ke dalam salah satu dari tiga cluster yang telah ditentukan. Centroid awal dapat dilihat pada Tabel 1.

Tabel 1. Centroid Awal

Cluster	1	2	3	4	5	6
C1	39	80	324	74	272	116
C2	2	8	3	1	2	4
C3	21	17	32	1	45	33

Pada tabel diatas terdapat 3 data yang diambil sebagai centroid awal dimana data penjualan dari bulan 1 sampai bulan 6 selanjutnya menghitung jarak data terhadap titik centroid awal dengan menggunakan Manhattan Distance, dalam setiap proses perhitungan jarak data ke masing-masing centroid untuk iterasi pertama dilakukan menggunakan data ke-1 sampai ke-5 sebagai contoh.

Perhitungan data C1 terhadap klaster

$$D1 = (58-39) + (191-80) + (89-324) + (84-74) + (120-272) + (32-116) = 611$$

$$D2 = (50-39) + (5-80) + (41-324) + (24-74) + (38-272) + (24-116) = 745$$

$$D3 = (16-39) + (37-80) + (46-324) + (3-74) + (45-272) + (42-116) = 716$$

Perhitungan data C2 terhadap klaster

$$D1 = (58-2) + (191-8) + (89-3) + (84-1) + (120-2) + (32-4) = 554$$

$$D2 = (50-2) + (5-8) + (41-3) + (24-1) + (38-2) + (24-4) = 168$$

$$D3 = (16-2) + (37-8) + (46-3) + (3-1) + (45-2) + (42-4) = 169$$

Perhitungan data C3 terhadap klaster

$$D1 = (58-21) + (191-17) + (89-32) + (84-1) + (120-45) + (32-33) = 427$$

$$D2 = (50-21) + (5-17) + (41-32) + (24-1) + (38-45) + (24-33) = 89$$

$$D3 = (16-21) + (37-17) + (46-32) + (3-1) + (45-45) + (42-33) = 50$$

Dari hasil perhitungan iterasi diatas dilakukan pengelompokan data, pada tahap ini data 1 sampai ke 50 akan dikelompokkan berdasarkan nilai minimal antara C1, C2 dan C3. Berikutnya menentukan centroid baru dengan menghitung mean dari proses iterasi, proses perhitungan sebagai berikut Perhitungan centroid baru pada C1.

$$\frac{30 + 60 + 866 + 68 + 39 + 136 + 0 + 72}{8} = 158.88$$

$$\frac{1212 + 524 + 220 + 100 + 80 + 256 + 738 + 548}{8} = 459.75$$

$$\frac{400 + 140 + 1596 + 190 + 324 + 156 + 401 + 364}{8} = 446.38$$

Perhitungan centroid baru pada C2

$$\frac{172 + 0 + 2 + 58 + 4 + 1 + 8 + 1 + 16 + 8 + 0 + 5 + 12 + 16 + 10 + 5 + 5 + 5 + 42 + 2 + 1 + 0 + 0 + 20}{24} = 16.38$$

$$\frac{61 + 0 + 0 + 52 + 0 + 2 + 0 + 3 + 2 + 1 + 0 + 5 + 20 + 11 + 38 + 15 + 15 + 7 + 0 + 8 + 3 + 2 + 2 + 0}{24} = 10.29$$

$$\frac{0 + 20 + 6 + 54 + 8 + 4 + 3 + 4 + 22 + 7 + 0 + 13 + 12 + 24 + 13 + 3 + 3 + 3 + 0 + 3 + 5 + 5 + 4 + 10}{24} = 9.42$$

Perhitungan centroid baru pada C3

$$\frac{58 + 50 + 16 + 206 + 4 + 0 + 24 + 0 + 16 + 0 + 24 + 24 + 2 + 16 + 0 + 10 + 21 + 20}{18} = 27.28$$

$$\frac{191 + 5 + 37 + 4 + 5 + 6 + 30 + 12 + 30 + 20 + 30 + 60 + 303 + 0 + 42 + 20 + 17 + 0}{18} = 45.11$$

$$\frac{89 + 41 + 46 + 88 + 39 + 59 + 40 + 38 + 16 + 44 + 14 + 61 + 8 + 92 + 6 + 2 + 32 + 10}{18} = 40.28$$

Hasil perhitungan nilai centroid baru untuk iterasi ke 2 seperti pada tabel berikut ini.

Tabel 2. Centroid Baru

Cluster	1	2	3	4	5	6
C1	158,88	459,75	446,38	204,5	344	241,75
C2	16,38	10,29	9,42	7,5	7,5	20,04
C3	27,28	45,11	40,28	31,89	81,89	34,56

C1 merupakan klaster 1, C2 merupakan klaster 2 dan C3 merupakan klaster 3, selanjutnya lakukan pengulangan hingga kondisi perhentian terpenuhi yaitu pada saat centroid awal dan centroid baru sudah tidak mengalami perubahan. Pada penelitian ini iterasi berhenti pada iterasi-5. Maka hasil akhir dapat dikelompokkan pada masing-masing klaster. Hasil proses hanya menampilkan 45 data dari 50 data seperti pada tabel berikut.

Tabel 3. Hasil Akhir Pengelompokan Sparepart Alat Berat

Data ke	Part Number	Jarak Terdekat	Klaster
1	1P 2299	432,63	Klaster 3
2	1R 0751	104,13	Klaster 2
3	2P 4004	86,67	Klaster 2
4	3F 5108	359,33	Klaster 2
5	3S 0336	214,90	Klaster 2
6	3S 9669	100,08	Klaster 2
7	4F 3657	156,74	Klaster 2
8	7H 3598	533,88	Klaster 3
9	8M 4987	94,67	Klaster 2
10	9S 4829	98,08	Klaster 2
11	14U-70-12520	99,08	Klaster 2
12	14X-30-13115	85,08	Klaster 2
13	14X-30-13142	105,08	Klaster 2
14	14X-30-13241	60,90	Klaster 2
15	14X-30-13310	87,08	Klaster 2
16	14X-32-11210	2014,33	Klaster 1
17	15B-LAMP	81,31	Klaster 2
18	20Y-27-11561	604,38	Klaster 3
19	20Y-30-11312	87,64	Klaster 2
20	20Y-32-05020	2291,00	Klaster 1
21	20Y-32-11340	572,88	Klaster 3
22	20Y-70-14520	65,05	Klaster 2
23	60X550	55,00	Klaster 2
24	65X550	71,90	Klaster 2
25	150-30-13430	41,36	Klaster 2
26	150-30-13442	48,33	Klaster 2
27	154-27-12320	484,88	Klaster 3
28	205-70-19570-STD	44,90	Klaster 2
29	205-70-19570-TGR	87,67	Klaster 2
30	205-70-19570-TGRUNIK	66,90	Klaster 2
31	600-181-4212	69,08	Klaster 2
32	600-181-4300	69,08	Klaster 2
33	600-185-2520	98,08	Klaster 2
34	600-211-1231	41,36	Klaster 2
35	600-311-8293	145,67	Klaster 2
36	01010-81655	467,85	Klaster 2
37	01010-81865	518,63	Klaster 3
38	01011-82465	122,23	Klaster 2
49	01643-31845	198,13	Klaster 2

40	01643-32060	524,13	Klaster 3
41	6150-11-5240	94,08	Klaster 2
42	6151-11-8610	95,08	Klaster 2
43	6151-62-1101	98,08	Klaster 2
44	07020-00000	117,26	Klaster 2
45	4185537	1303,33	Klaster 1

Sehingga hasil dari perhitungan dimana klaster 1 adalah Paling Laris dengan 3 data yaitu data ke (16, 20, 45), klaster 2 adalah Laris dengan 39 data yaitu data ke (2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 17, 19, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 39, 41, 42, 43, 44, 47, 48, 49, 50) dan klaster 3 adalah Tidak Laris dengan 8 data yaitu data ke (1, 8, 18, 21, 27, 37, 40, 46).

Kategori Paling Laris meliputi suku cadang yang paling sering dibeli oleh konsumen, sehingga perusahaan harus memastikan ketersediaan stok yang cukup untuk kategori ini. Kategori Laris mencakup suku cadang yang memiliki permintaan cukup tinggi, tetapi tidak sekuat kategori pertama. Sementara itu, suku cadang dalam kategori Tidak Laris adalah suku cadang dengan permintaan yang rendah, yang bisa dijadikan pertimbangan untuk mengurangi stok atau bahkan dihentikan pengadaannya.

IV. KESIMPULAN

Metode K-Means Clustering telah terbukti efektif untuk digunakan dalam analisis penjualan suku cadang alat berat di PT. Patria Jaya Mandiri. Dengan menggunakan metode ini, suku cadang dapat dikelompokkan ke dalam tiga kategori berdasarkan tingkat permintaan: Paling Laris, Laris, dan Tidak Laris. Dalam penelitian ini, proses iterasi K-Means mencapai kondisi optimal pada iterasi ke-5, di mana tidak ada lagi perubahan signifikan pada centroid yang menunjukkan bahwa hasil clustering telah stabil. Hasil akhirnya menunjukkan bahwa 3 suku cadang termasuk dalam kategori Paling Laris, 39 suku cadang termasuk dalam kategori Laris, dan 8 suku cadang dikategorikan sebagai Tidak Laris.

Dengan adanya hasil clustering ini, PT. Patria Jaya Mandiri dapat menggunakan informasi tersebut untuk mengoptimalkan pengelolaan stok suku cadang mereka. Dengan cara ini, perusahaan dapat memaksimalkan keuntungan dengan memastikan ketersediaan suku cadang yang paling dibutuhkan konsumen, sekaligus mengurangi potensi kerugian dari suku cadang yang kurang diminati.

REFERENSI

- [1] L. Yana Siregar, M. Irwan Padli Nasution Prodi Manajemen, and U. Negeri Islam Sumatera Utara, "Perkembangan Teknologi Informasi Terhadap Peningkatan Bisnis Online," *HIRARKI: Jurnal Ilmiah Manajemen dan Bisnis*, vol. 2, no. 1, pp. 71–75, 2020, doi: 10.30606/hjimb.
- [2] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *Jurnal Teknologi dan Sistem Informasi (JTSI)*, vol. 2,

- no. 2, p. 100, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSTI>
- [3] R. Kesuma Dinata, N. Hasdyna, and N. Azizah, "Analisis K-Means Clustering pada Data Sepeda Motor," 2020.
- [4] T. Hardiani, "Analisis Clustering Kasus Covid 19 di Indonesia Menggunakan Algoritma K-Means," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 11, no. 2, pp. 156–165, Aug. 2022, doi: 10.23887/janapati.v11i2.45376.
- [5] P. Marpaung, I. Pebrian, and W. Putri, "Penerapan Data Mining Untuk Pengelompokan Kepadatan Penduduk Kabupaten Deli Serdang Menggunakan Algoritma K-Means," *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, vol. 6, no. 2, pp. 64–70, 2023.
- [6] E. Muningsih, I. Maryani, and V. R. Handayani, "Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa," *Jurnal Sains dan Manajemen*, vol. 9, no. 1, 2021, [Online]. Available: www.bps.go.id
- [7] I. Nuryani and D. Darwis, "Analisis Clustering Pada Pengguna Brand Hp Menggunakan Metode K-Means", *SNASIKOM*, vol. 1, no. 1, pp. 190–211, Jul. 2021.
-