

## Penerapan Algoritma K-Means untuk Pengelompokan Pola Penjualan Oleh-Oleh Khas Medan

Aldaka Triyanti Agustina Basuki<sup>1</sup>, Zahra Puspita Paramastri<sup>2</sup>, Dina Fitri Rachmayanti<sup>3</sup>,  
Dewi Aminah Chan<sup>4</sup>, Dasril Aldo<sup>5</sup>

*Department of Informatics, Telkom University  
Purwokerto Kidul 53147, Indonesia*

<sup>1</sup>aldakatriyanti2208@gmail.com

<sup>2</sup>zahrarama10@gmail.com

<sup>3</sup>dinafitri.ry@gmail.com

<sup>4</sup>dewiaminahchan@gmail.com

<sup>5</sup>dasril@ittelkom-pwt.ac.id

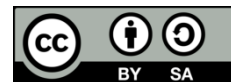
Dikirim pada 20-11-2024, Direvisi pada 27-11-2024, Diterima pada 04-12-2024

### Abstrak

Penelitian ini bertujuan untuk menerapkan algoritma *K-Means* dalam pengelompokan pola penjualan oleh-oleh khas Medan guna memberikan wawasan strategis bagi pelaku usaha dalam manajemen stok dan pemasaran. Data yang digunakan meliputi stok awal dan jumlah penjualan produk oleh-oleh khas Medan, seperti Bika Ambon, Kopi Sidikalang, Dodol Durian, Teri Medan, dll. Data diolah menggunakan metode *Elbow* untuk menentukan jumlah cluster optimal, sementara validasi dilakukan dengan *Silhouette Score* dan menghasilkan nilai 0.75 yang menandakan clustering yang baik. Hasil analisis menghasilkan tiga cluster: produk dengan stok dan penjualan moderat, produk dengan permintaan tinggi, serta produk dengan stok berlebih tetapi penjualan rendah. Temuan ini menunjukkan bahwa algoritma *K-Means* mampu mengidentifikasi pola penjualan secara sistematis, membantu pelaku usaha dalam mengambil keputusan berbasis data, seperti penyesuaian stok, strategi promosi, dan pengurangan pemborosan. Penelitian ini diharapkan dapat berkontribusi pada pengembangan strategi bisnis yang lebih efektif yang akan bermanfaat bagi pelaku usaha lokal dan komunitas pemasaran dalam sektor pariwisata lokal, sehingga meningkatkan daya saing produk oleh-oleh khas Medan di pasar yang kompetitif.

**Kata Kunci:** Algoritma K-Means, Pengelompokan, Pola Penjualan, Oleh-Oleh Khas Medan, Strategi Pemasaran

*Ini adalah artikel akses terbuka di bawah lisensi [CC BY-SA](#).*



---

### Penulis Koresponden:

Aldaka Triyanti Agustina Basuki

Program Studi Teknik Informatika, Universitas Telkom Kampus Purwokerto, Jl. D.I Panjaitan No.128 Purwokerto, 53147 Kabupaten Banyumas, Jawa Tengah, Indonesia Email: aldakatriyanti2208@gmail.com

---

### I. PENDAHULUAN

Medan, sebagai salah satu kota di Indonesia yang kaya akan budaya dan kuliner, memiliki beragam produk oleh-oleh yang diminati oleh wisatawan. Oleh-oleh khas Medan tidak hanya mencerminkan keunikan budaya lokal, tetapi juga berkontribusi pada perekonomian daerah. Dengan meningkatnya persaingan di sektor ini, penting bagi pelaku usaha untuk memahami pola penjualan mereka agar dapat meningkatkan efisiensi dan efektivitas dalam manajemen stok dan pemasaran. Penelitian sebelumnya menunjukkan bahwa pengelompokan data penjualan menggunakan K-Means dapat membantu dalam mengidentifikasi produk yang paling laris, serta produk yang kurang diminati, sehingga pelaku usaha dapat mengambil keputusan yang lebih baik dalam pengelolaan produk [1], [2].

Terdapat produk dengan stok dan penjualan yang cukup banyak, tetapi penjualannya masih relatif rendah, yang menandakan adanya permintaan yang stabil. Adapun produk yang permintaannya relatif

tinggi, yang menandakan produk tersebut sangat diminati oleh pelanggan. Sementara itu, terdapat produk yang memiliki stok yang melimpah namun memerlukan strategi pemasaran untuk meningkatkan penjualannya. Pengelompokan data merupakan teknik analisis yang penting dalam berbagai bidang, termasuk bisnis dan pemasaran. Salah satu metode yang banyak digunakan untuk pengelompokan adalah algoritma K-Means [3], [4]. Algoritma ini berfungsi untuk mengelompokkan data berdasarkan kesamaan karakteristik, sehingga dapat membantu dalam mengidentifikasi pola dan trend dalam data yang ada [5]. Dalam konteks penjualan oleh-oleh khas Medan, penerapan algoritma K-Means dapat memberikan wawasan yang berharga bagi pelaku usaha dalam memahami perilaku konsumen dan pola penjualan yang terjadi.

Salah satu studi yang relevan adalah penelitian oleh Gustriansyah et al. yang menggunakan pendekatan K-Means untuk mengelompokkan data pelanggan berdasarkan analisis RFM (*Recency, Frequency, Monetary*). Penelitian ini menunjukkan bahwa pengelompokan dapat membantu perusahaan dalam mengelola produk yang kurang diminati dan memberikan penawaran yang lebih baik untuk produk yang tidak terjual [6], [7]. Hal ini menunjukkan bahwa pendekatan serupa dapat diterapkan pada penjualan oleh-oleh khas Medan untuk meningkatkan pemahaman tentang preferensi konsumen dan perilaku pembelian. Lebih lanjut, penelitian oleh Tabianan et al. menunjukkan bahwa penggunaan algoritma K-Means dalam analisis cluster dapat membantu dalam menentukan strategi promosi produk. Penelitian ini menekankan pentingnya pengelompokan dalam merumuskan strategi pemasaran yang lebih efektif berdasarkan karakteristik produk dan perilaku konsumen [8]. Dengan memanfaatkan data penjualan yang ada, pelaku usaha dapat mengidentifikasi tren dan pola yang dapat digunakan untuk merumuskan strategi pemasaran yang lebih efektif. Dalam konteks yang lebih luas, algoritma K-Means juga telah diterapkan dalam berbagai penelitian untuk meningkatkan kinerja penjualan di sektor lain.

Sebagai contoh, penelitian oleh Oyewole dan Thopil menunjukkan bahwa pengelompokan data penjualan dapat membantu dalam menentukan strategi pemasaran yang paling tepat untuk produk tertentu [9]. Penelitian ini menekankan pentingnya pengelompokan dalam mengoptimalkan manajemen penjualan, yang dapat diterapkan pada penjualan oleh-oleh khas Medan. Kelebihan dari algoritma K-Means adalah kemampuannya dalam menangani data besar dan kompleks, serta kemudahan dalam implementasinya. Penelitian oleh Brahmana et al. menunjukkan bahwa pengelompokan menggunakan K-Means dapat meningkatkan akurasi model peramalan dengan mengelompokkan data berdasarkan pola penjualan yang ada [10]. Dengan demikian, penerapan K-Means dalam pengelompokan pola penjualan oleh-oleh khas Medan tidak hanya relevan tetapi juga dapat memberikan kontribusi signifikan terhadap pengembangan strategi bisnis yang lebih baik.

Dalam penelitian ini, bertujuan untuk menerapkan algoritma K-Means dalam pengelompokan pola penjualan oleh-oleh khas Medan. Dengan menggunakan data penjualan yang dikumpulkan dari berbagai sumber, kami berharap dapat mengidentifikasi pola dan tren yang ada, serta memberikan rekomendasi yang berguna bagi pelaku usaha dalam meningkatkan kinerja penjualan mereka. Penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan ekonomi lokal dan meningkatkan daya saing produk oleh-oleh khas Medan di pasar yang semakin kompetitif.

## II. METODE PENELITIAN

### A. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan algoritma *K-Means* dalam pengelompokan produk oleh-oleh khas Medan berdasarkan pola penjualannya. Tujuan utama dari penelitian ini adalah mengidentifikasi kelompok produk dengan pola penjualan dan ketersediaan stok yang serupa, sehingga dapat mendukung keputusan strategis dalam manajemen inventaris dan perencanaan pemasaran. Analisis dilakukan menggunakan data historis yang tersedia dalam format CSV, yang berisi informasi seperti ID Produk, Nama Produk, Kemasan, Stok Awal, dan Jumlah Terjual.

### B. Data Penelitian

Data yang digunakan dalam penelitian ini diperoleh dengan cara mengunjungi salah-satu toko yang menjual produk oleh-oleh khas Medan yang berada di bandara kualanamu. Dataset yang digunakan berisi informasi mengenai produk oleh-oleh khas Medan dan disusun untuk mencerminkan pola penjualan aktual. Adapun kolom-kolom dalam dataset ini adalah sebagai berikut:

1. ID Produk: Kolom ini berisi identitas unik yang diberikan pada setiap produk oleh-oleh untuk membedakannya satu sama lain.
2. Nama Produk: Kolom ini berisi deskripsi atau nama produk oleh-oleh khas Medan, seperti Bika Ambon, Kopi Sidikalang, atau Sirup Markisa, yang berfungsi untuk mengidentifikasi jenis produk.

3. Kemasan: Kolom ini menunjukkan tipe kemasan produk oleh-oleh, misalnya dalam bentuk kotak, bungkus, atau botol, yang membantu dalam memahami variasi penyajian produk di pasaran.
4. Stok Awal: Kolom ini berisi jumlah stok produk yang tersedia di awal periode analisis. Data ini penting untuk melihat persediaan awal yang ada sebelum penjualan dilakukan.
5. Terjual: Kolom ini menunjukkan jumlah produk yang terjual selama periode analisis. Data ini akan menjadi acuan utama dalam memahami pola penjualan masing-masing produk. Sebagai basis data, data ini dianalisis menggunakan metode *K-Means* untuk menemukan pola penjualan dan ketersediaan produk oleh-oleh di Medan. Ini dapat membantu strategi pemasaran dan perencanaan inventaris menjadi lebih efisien.

### C. Prosedur Penelitian

#### 1. Pengumpulan dan Pra-pemrosesan Data

Pengumpulan data dilakukan dengan cara mengunjungi salah-satu toko yang menjual produk oleh-oleh khas medan yang berada di bandara kualanamu. Pra-pemrosesan data dilakukan untuk memastikan data siap untuk diproses. Dataset diimpor dari file CSV, karena CSV adalah format yang mudah diakses untuk menyimpan data tabular. Kemudian dilakukan pemeriksaan data agar tidak terdapat nilai kosong, karena akan mempengaruhi data saat diproses.

#### 2. Analisis Pendahuluan

Statistik deskriptif dilakukan untuk memahami distribusi stok dan penjualan. *Scatter plot* dibuat untuk mengidentifikasi pola awal dalam data.

#### 3. Clustering dengan *K-Means*

Jumlah *cluster* ( $k$ ) ditentukan menggunakan metode Elbow, dengan melihat grafik *Within-Cluster Sum of Squares* (WCSS) untuk menemukan titik optimal. Algoritma *K-Means* diterapkan untuk mengelompokkan produk berdasarkan stok awal dan jumlah produk yang terjual.

#### 4. Evaluasi dan Validasi Hasil Clustering

Validasi hasil clustering dilakukan menggunakan Silhouette Score untuk mengukur efektivitas pengelompokan. Interpretasi visual dilakukan melalui scatter plot untuk memahami distribusi cluster.

### D. Proses Clustering

Langkah-langkah algoritma K-Means yang digunakan: Menentukan jumlah cluster ( $k$ ) berdasarkan hasil metode Elbow. Menginisialisasi centroid secara acak. Mengelompokkan data ke centroid terdekat berdasarkan jarak Euclidean. Memperbarui posisi centroid berdasarkan rata-rata data dalam tiap cluster. Mengulangi proses hingga centroid stabil atau iterasi maksimum tercapai.

### E. Evaluasi dan Analisis Data

Setelah clustering dilakukan cluster yang memiliki stok rendah dan tingkat penjualan tinggi diidentifikasi sebagai produk yang memerlukan restock cepat. Cluster dengan stok tinggi namun tingkat penjualan rendah perlu meningkatkan strategi penjualan atau mengurangi jumlah produksi sampai stok dan penjualan seimbang dengan permintaan pelanggan.

## III. HASIL DAN PEMBAHASAN

Dalam penelitian ini, dilakukan analisis pola penjualan produk oleh-oleh khas Medan menggunakan metode clustering berbasis *K-Means*. Dataset yang digunakan berisi atribut penting seperti stok awal dan jumlah terjual untuk setiap produk. Proses analisis menghasilkan tiga *cluster* utama yang mengelompokkan produk berdasarkan pola penjualan.

Pada cluster pertama mencakup produk dengan stok dan penjualan moderat yang banyak tetapi penjualan rendah, yang menunjukkan permintaan yang stabil tetapi tidak terlalu tinggi. Cluster kedua terdiri dari produk dengan permintaan tinggi, dimana hampir seluruh stok terjual, yang menunjukkan bahwa pelanggan sangat menyukai produk ini. Cluster ketiga terdiri dari produk dengan stok yang banyak tetapi penjualan rendah, yang menunjukkan bahwa produk ini mungkin membutuhkan lebih banyak strategi pemasaran untuk meningkatkan penjualan.

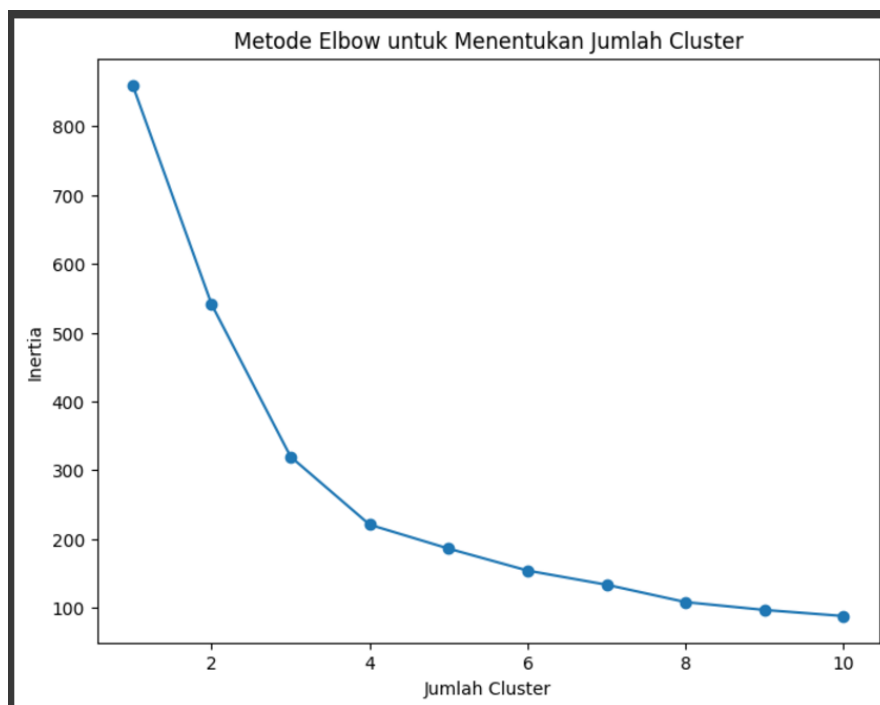
```

Jumlah total data: 430
  ID Produk  Nama Produk  Cluster
0           1      Bika Ambon    0
1           2      Bolu Meranti   0
2           3      Dodol Durian   1
3           4      Teri Medan     1
4           5      Pancake Durian  0
..          ...      ...           ...
425         426     Kue Nagasari   1
426         427     Kue Talam Pandan 0
427         428     Dendeng Sapi   2
428         429     Tape Singkong  2
429         430     Dodol Kopi     2

[430 rows x 3 columns]

```

Gambar 1. Data yang akan digunakan



Gambar 2. Grafik Elbow

Grafik Elbow Method menunjukkan hubungan antara jumlah *cluster* ( $k$ ) dan inertia. Optimalnya adalah jumlah cluster di titik di mana penurunan inertia mulai melambat, biasanya disebut "elbow". Berdasarkan grafik ini, dipilih jumlah *cluster* optimal sebanyak tiga ( $k=3$ ). Ketiga *cluster* tersebut memiliki karakteristik sebagai berikut:

1. Cluster 0: Produk dengan rata-rata stok awal 531 unit dan rata-rata penjualan 367 unit. *Cluster* ini mencakup produk dengan stok sedang dan penjualan yang juga moderat.
2. Cluster 1: Produk dengan rata-rata stok awal 1089 unit dan rata-rata penjualan 944 unit. *Cluster* ini menunjukkan produk dengan permintaan tinggi, dimana hampir seluruh stok terjual.
3. Cluster 2: Produk dengan rata-rata stok awal 1171 unit, tetapi rata-rata penjualan hanya 342 unit. *Cluster* ini menunjukkan produk dengan stok tinggi namun penjualan rendah.

```

Pengkategorian Produk berdasarkan Cluster:
Cluster  Jumlah Produk
0         0           142
1         1           132
2         2           156

Total Produk di Seluruh Cluster (Hasil Penjumlahan Cluster 0, 1, dan 2): 430

```

Gambar 3. Pengkategorian Cluster

Produk dalam kelompok cluster 0 memiliki tingkat penjualan sedang. Mereka tidak laku keras, tetapi tetap memiliki pembeli reguler. Manajemen perlu memastikan stok mencukupi, tetapi tidak terlalu berlebihan.

```

Cluster 0:
ID Produk      Nama Produk      Cluster
0             1             Bika Ambon      0
1             2             Bolu Meranti    0
4             5             Pancake Durian  0
6             7             Kacang Sihobuk 0
9             10            Sirup Markisa Medan 0
..           ...           ...             ...
414           415            Gula Aren Batak 0
415           416            Kue Kering Batak 0
416           417            Kue Gandus      0
422           423            Gulai Ikan Mas  0
426           427            Kue Talam Pandan 0

[142 rows x 3 columns]

```

Gambar 4. Produk pada cluster 0

Produk ini sangat diminati, sehingga membutuhkan pengelolaan stok yang cepat. *Restock* harus dilakukan secara rutin untuk menghindari kekosongan barang di rak.

```

Cluster 1:
ID Produk      Nama Produk      Cluster
2             3             Dodol Durian    1
3             4             Teri Medan      1
5             6             Kopi Sidikalang 1
12            13            Roti Ketawa     1
15            16            Sale Pisang     1
..           ...           ...             ...
413           414            Manisan Pala    1
418           419            Manisan Kedondong 1
419           420            Bubuk Jahe Merah 1
424           425            Lapek Bugis    1
425           426            Kue Nagasari   1

[132 rows x 3 columns]

```

Gambar 5. Produk pada cluster 1

*Cluster* ini menunjukkan produk dengan stok besar tetapi jarang dibeli. Manajemen perlu mempertimbangkan strategi promosi atau pengurangan stok agar tidak terlalu banyak barang menumpuk.

```

Cluster 2:
  ID Produk  Nama Produk  Cluster
7           8   Manisan Jambu  2
8           9   Lempok Durian  2
10          11   Sambal Tuktuk  2
13          14   Ikan Sale      2
14          15   Keripik Sanjai  2
..          ...   ...            ..
421         422   Tepung Singkong  2
423         424   Kue Bawang     2
427         428   Dendeng Sapi   2
428         429   Tape Singkong  2
429         430   Dodol Kopi     2

[156 rows x 3 columns]

```

Gambar 6. Gambar produk pada cluster 2

Hasil clustering menunjukkan pola distribusi penjualan yang beragam, membantu dalam mengidentifikasi strategi pemasaran dan manajemen stok: Cluster 0 membutuhkan pengelolaan stok yang seimbang, mengingat sebagian besar produknya memiliki tingkat penjualan moderat. Cluster 1 dapat menjadi prioritas untuk meningkatkan stok karena menunjukkan permintaan yang konsisten. Cluster 2 memerlukan evaluasi lebih lanjut, misalnya dengan menawarkan diskon atau promosi guna meningkatkan penjualan.

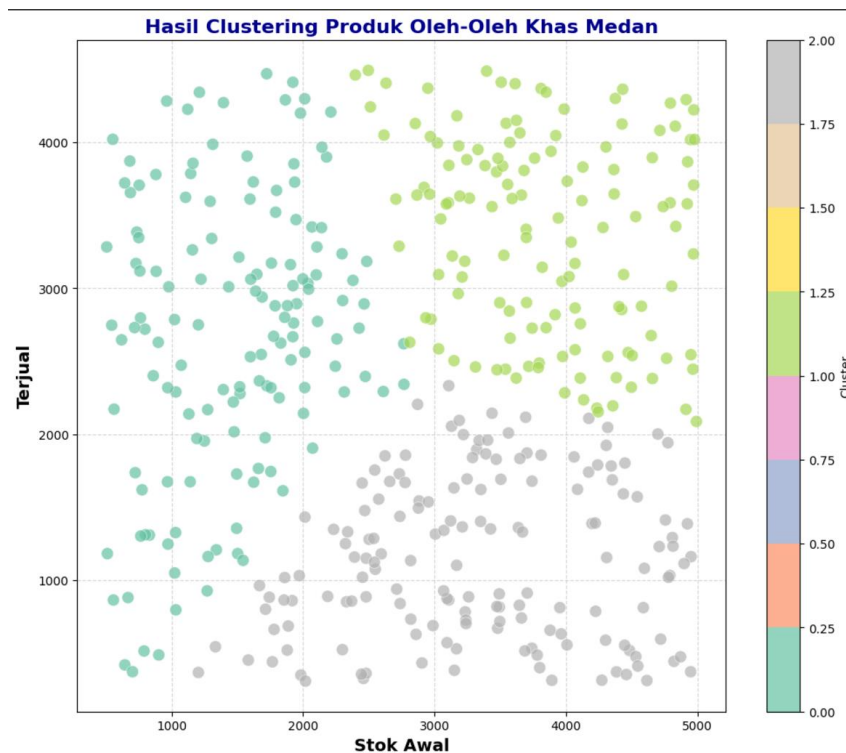


Fig. 7. Gambar Visualisasi Cluster dalam grafik

#### IV. KESIMPULAN

Penelitian ini menunjukkan bahwa penerapan algoritma K-Means dapat membantu manajemen memahami pola penjualan produk ATK dengan cara yang sistematis. Dengan membagi produk ke dalam tiga kategori (stok dan penjualan moderat, stok tinggi dengan penjualan tinggi, dan stok tinggi dengan penjualan rendah), manajemen dapat mengambil keputusan berbasis data yang lebih tepat. Produk dalam cluster dengan tingkat penjualan tinggi membutuhkan restock yang cepat dan teratur untuk menghindari

kekosongan stok yang dapat mengganggu kepuasan pelanggan. Produk dengan stok besar tetapi penjualan rendah memerlukan strategi promosi yang efektif, seperti diskon atau kampanye pemasaran, untuk meningkatkan daya tarik pembeli. Dengan mengetahui produk yang memiliki potensi stok berlebih, manajemen dapat mengurangi pemborosan sumber daya seperti ruang penyimpanan dan modal yang terkunci. Selain itu, penelitian ini dapat diperluas dengan menambahkan atribut lain, seperti harga produk, tren musiman, atau lokasi penjualan, untuk mendapatkan hasil clustering yang lebih akurat dan relevan dalam mendukung pengambilan keputusan manajemen.

#### UCAPAN TERIMAKASIH

Terima kasih kami panjatkan kepada Allah Yang Maha Kuasa atas terselesainya penelitian "Penerapan Clustering K-Means Untuk Pengelompokan Pola Penjualan Oleh-Oleh Khas Medan" dengan baik. Penelitian ini merupakan kontribusi dalam pengembangan kecerdasan artifisial untuk strategi pemasaran produk pariwisata di Indonesia. Kami menyadari masih terdapat keterbatasan, sehingga kritik dan saran konstruktif dari para pembaca sangat diharapkan untuk penyempurnaan penelitian selanjutnya.

#### DAFTAR PUSTAKA

- [1] P. Edastama, A. S. Bist, and A. Prambudi, 'Implementation Of Data Mining On Glasses Sales Using The Apriori Algorithm', *Int. J. Cyber IT Serv. Manag.*, vol. 1, no. 2, pp. 159–172, Oct. 2021.
- [2] U. Radtke, 'K-means cluster analysis of hourly measured power demand in the district heating system in Kaposvár : Putting theoretical figures into practical use', *Reg. Bus. Stud.*, vol. 14, no. 1, pp. 57–72, Jun. 2022.
- [3] M. Annas and S. N. Wahab, 'Data Mining Methods: K-Means Clustering Algorithms', *Int. J. Cyber IT Serv. Manag.*, vol. 3, no. 1, pp. 40–47, Mar. 2023.
- [4] F. Li, Y. Qian, and J. Wang, 'GoT: a Growing Tree Model for Clustering Ensemble', *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 9, pp. 8349–8356, May 2021.
- [5] K. Tabianan, S. Velu, and V. Ravi, 'K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data', *Sustainability*, vol. 14, no. 12, p. 7243, Jun. 2022.
- [6] R. Gustriansyah, N. Suhandi, and F. Antony, 'Clustering optimization in RFM analysis Based on k-Means', *Indones. J. Electr. Eng. Comput. Sci.*, vol. 18, no. 1, p. 470, Apr. 2020.
- [7] A. Adli, F. Prastyasari, D. Handani, and K. Artana, 'LNG Distribution Optimization using Set Partitioning Problem Method', *IOP Conf. Ser. Earth Environ. Sci.*, vol. 972, no. 1, p. 012082, Jan. 2022.
- [8] O. N. Oyelade and A. E. Ezugwu, 'A case-based reasoning framework for early detection and diagnosis of novel coronavirus', *Inform. Med. Unlocked*, vol. 20, p. 100395, 2020.
- [9] S. Mukhi and B. E. Medise, 'Faktor yang Memengaruhi Penurunan Cakupan Imunisasi pada Masa Pandemi Covid-19 di Jakarta', *Sari Pediatri*, vol. 22, no. 6, p. 336, Apr. 2021.
- [10] U. F. Kurniawati, 'Dampak Perubahan Penggunaan Lahan Terhadap Besaran Stok Karbon di Kota Surabaya', *J. Penataan Ruang*, vol. 16, no. 1, p. 54, Mar. 2021
- [11] Erdi Amos Saputra, Yessica Nataliani, 'Analisis Pengelompokan Data Nilai Siswa untuk Menentukan Siswa Berprestasi Menggunakan Metode Clustering K-Means,' *Jurnal of Information Systems and Informatics.*, vol. 3, no. 3, 2021.
- [12] Tifani Amalina, Danendra Bima Adhi Pramana, Betha Nurina Sari, 'Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food,' *Jurnal Ilmiah Wahana Pendidikan.*, vol 8, no 15, pp. 574-584, Sep. 2022.
- [13] Elin Mayoana Fitri, Ryan Randy Suryon, and Agus Wantoro, 'Klasterisasi Data Penjualan Berdasarkan Wilayah Menggunakan Metode K-Means pada PT XYZ,' *Jurnal Komputasi.*, vol 11, no. 2, Nov 2023.
- [14] Muhammad Rafi Nahjan, Nono Heryana, and Apriade Voutama, 'Implementasi Rapidminer dengan Metode Clustering K-Means untuk Analisa Penjualan pada Toko OJ CELL,' *Jurnal Mahasiswa Teknik Informatika.*, vol 7, no 1, Feb 2023.
- [15] Nur Afiasari, Nana Suarna, and Nining Rahaningsih, 'Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means,' *Jurnal SAINTEKOM.*, vol 13, no 1, Mar 2023.