

## CLUSTERING DATA INDONESIAN FOOD DELIVERY MENGGUNAKAN METODE K-MEANS PADA GOFOOD PRODUCT LIST

Muhamad Farhan Nugraha<sup>1</sup>, Martanto<sup>2</sup>, Umi Hayati<sup>3</sup>

<sup>1,3</sup>Teknik Informatika, STMIK IKMI Cirebon

<sup>2</sup>Manajemen Informatika, STMIK IKMI Cirebon

Jl. Perjuangan 10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135

*m.farhan.nugraha@gmail.com*

### ABSTRAK

Dalam era transformasi digital yang berkembang pesat, layanan pengiriman makanan, seperti GoFood di Indonesia, menunjukkan dampak konkret pada perubahan perilaku konsumen. Penelitian ini bertujuan untuk memahami bagaimana transformasi digital memengaruhi interaksi konsumen dengan layanan pengiriman makanan, khususnya melalui GoFood, serta implikasinya dalam bisnis dan pengalaman konsumen di Indonesia. Metode *k-means clustering* digunakan untuk mengelompokkan produk makanan ke dalam kelompok yang serupa berdasarkan atribut *category*, mengidentifikasi kesamaan dalam kelompok produk yang mungkin mencerminkan kecenderungan konsumen yang serupa. Melalui penelitian ini, akan terungkap informasi yang bernilai tentang keberagaman preferensi konsumen, memungkinkan pelaku industri untuk menyesuaikan strategi pemasaran mereka sesuai dengan permintaan pasar yang berubah. Dengan demikian, penelitian ini tidak hanya dapat memberikan kontribusi pada pemahaman akademis tetapi juga memiliki potensi dampak positif secara praktis dalam mendukung pertumbuhan dan peningkatan kualitas layanan di sektor pengiriman makanan di Indonesia. Hasil pengelompokan menggunakan algoritma *K-Means Clustering* diperoleh terbaik sebanyak 8 *cluster* dengan nilai DBi= -0.407, yaitu dengan anggota *cluster* 0=7763 *items*, *cluster* 1=1 *items*, *cluster* 2=1 *items*, *cluster* 3=1 *items*, *cluster* 4=1 *items*, *cluster* 5=1 *items*, *cluster* 6=1 *items* dan *cluster* 7=1 *items* yang artinya adalah melihat karakteristik konsumen dengan tujuan untuk meningkatkan penjualan.

**Kata kunci:** analisis penjualan, K-Means, dataset transaksi, produk snack, minuman

### 1. PENDAHULUAN

Dalam era pesatnya perkembangan teknologi informasi, aspek informatika telah menjadi pilar utama yang membentuk dinamika berbagai sektor kehidupan. Perkembangan ini tidak hanya memengaruhi teknologi itu sendiri, tetapi juga merasuk ke dalam *domain* bisnis, pendidikan, dan beragam aspek kehidupan sehari-hari. Kehadiran layanan pengiriman makanan, khususnya melalui *platform* GoFood di Indonesia, menjadi salah satu contoh konkret bagaimana transformasi digital mengubah cara kita berinteraksi dengan layanan konsumen. Dalam konteks ini, analisis data menjadi semakin relevan untuk memahami pola dan kecenderungan di balik preferensi konsumen, memungkinkan penyedia layanan dan restoran untuk menyelaraskan penawaran mereka dengan tuntutan pasar yang terus berubah. Penelitian ini fokus pada eksplorasi analisis data menggunakan metode *k-means clustering* pada daftar produk GoFood, dengan harapan dapat memberikan kontribusi pada pengembangan strategi pemasaran yang lebih cerdas dan pemahaman yang lebih mendalam tentang preferensi konsumen di industri pengiriman makanan di Indonesia.

Namun, seiring dengan kemajuan pesat dalam layanan pengiriman makanan, terutama melalui *platform* GoFood, muncul sejumlah permasalahan dan tantangan yang perlu diatasi. Pertama, volume data yang terus meningkat dari transaksi dan preferensi konsumen menimbulkan kompleksitas analisis, membutuhkan metode yang canggih untuk

mengungkap pola yang bermakna. Kedua, variasi preferensi konsumen yang beragam di berbagai kota di Indonesia menambah kompleksitas analisis, dengan kebutuhan untuk memahami dinamika pasar lokal. Ketiga, umpan balik konsumen terhadap *platform* GoFood sangat beragam. Selain itu, tingginya persaingan di industri pengiriman makanan menuntut strategi pemasaran yang cerdas dan diferensiasi produk yang lebih baik. Ketidakpastian dalam tren preferensi konsumen juga menjadi tantangan serius bagi penyedia layanan dan restoran dalam merancang menu yang sesuai dengan selera pasar yang terus berubah. Oleh karena itu, penelitian ini berusaha untuk mengatasi permasalahan-permasalahan tersebut melalui analisis data menggunakan metode *k-means clustering*, memberikan pandangan yang lebih mendalam tentang pola preferensi konsumen, dan menyusun rekomendasi strategis untuk meningkatkan daya tarik produk di tengah persaingan yang semakin ketat. Dengan mengumpulkan dan menganalisis data konsumen, restoran dapat mengidentifikasi pola dan tren yang muncul, memungkinkan mereka untuk merancang menu yang lebih responsif terhadap preferensi pelanggan. Hasil dari analisis *k-means clustering* dapat membantu restoran memahami preferensi utama pelanggan, memisahkan item-menu yang paling diminati, dan mengadaptasi penawaran makanan mereka sesuai dengan segmen pasar yang berbeda. Dengan demikian, restoran dapat secara proaktif menyesuaikan strategi mereka dengan perubahan selera pasar, meningkatkan daya saing

mereka, dan memastikan kepuasan pelanggan yang berkelanjutan.

Sejumlah studi sebelumnya telah membahas aspek-aspek terkait dalam industri pengiriman makanan dan preferensi konsumen. Sebagai contoh, penelitian oleh [1] fokus pada analisis preferensi konsumen di *platform* pengiriman makanan namun belum menggunakan metode *k-means clustering* untuk memahami kelompok-kelompok produk yang mungkin memiliki karakteristik konsumen yang serupa. Di sisi lain, [2] seiring berkembangnya dunia perdagangan elektronik, hal itu diikuti dengan kian meningkatnya aktivitas jual beli secara *online*. Selain itu, sebuah penelitian oleh [3] berhasil memberikan wawasan mendalam tentang faktor-faktor yang memengaruhi keputusan konsumen dalam memilih layanan pengiriman makanan. Namun, keterbatasan penelitian ini terletak pada kurangnya penerapan metode klustering untuk memahami secara lebih terperinci bagaimana kelompok produk makanan tertentu dapat diidentifikasi berdasarkan karakteristik konsumen. Melalui evaluasi literatur ini, tampaknya masih ada kekosongan pengetahuan yang dapat diisi oleh penelitian ini. Dengan menerapkan metode *k-means clustering* pada daftar produk GoFood, penelitian ini diharapkan dapat memperluas pemahaman tentang kelompok-kelompok produk dengan karakteristik konsumen yang serupa, mengisi kesenjangan pengetahuan yang ada, dan memberikan kontribusi pada literatur yang ada dalam bidang ini.

Penelitian ini bertujuan untuk mengaplikasikan metode *k-means clustering* pada data produk GoFood di Indonesia dengan tujuan utama mengidentifikasi kelompok-kelompok produk yang memiliki karakteristik konsumen serupa. Dengan demikian, penelitian ini bertujuan untuk memberikan pemahaman yang lebih mendalam tentang preferensi konsumen dalam konteks layanan pengiriman makanan. Signifikansi penelitian ini terletak pada kemampuannya untuk mengisi kesenjangan pengetahuan yang terdapat dalam literatur sebelumnya, khususnya dalam menganalisis dampak variasi geografis pada preferensi konsumen menggunakan metode klustering. Dengan memahami pola preferensi ini, penelitian ini dapat memberikan wawasan berharga bagi penyedia layanan dan restoran untuk merancang strategi pemasaran yang lebih efektif dan menyelaraskan penawaran mereka dengan kebutuhan pasar yang bervariasi. Secara lebih luas, penelitian ini diharapkan memberikan kontribusi pada literatur dan pemahaman kita tentang interaksi antara teknologi informasi dan konsumen, khususnya dalam konteks industri pengiriman makanan yang semakin berkembang di Indonesia. Potensi manfaat praktis dari hasil penelitian ini mencakup kemungkinan meningkatkan daya saing bisnis, memandu pengembangan menu yang lebih sesuai dengan preferensi konsumen, dan meningkatkan kepuasan pelanggan secara keseluruhan.

Dalam merespon tantangan dan tujuan penelitian, pendekatan metodologi yang akan digunakan melibatkan penerapan metode *k-means clustering* pada data produk GoFood di Indonesia. Data ini mencakup informasi seperti nama menu, harga, dan kategori makanan dari berbagai kota di Indonesia. Setelah mengumpulkan data, tahap pembersihan dan normalisasi data akan dilakukan untuk mempersiapkannya untuk analisis. Metode *k-means clustering* akan diterapkan untuk mengelompokkan produk makanan ke dalam kelompok dengan karakteristik konsumen serupa, berdasarkan atribut seperti jenis makanan dan harga. Dengan menggunakan metode ini, penelitian ini diharapkan memberikan wawasan lebih mendalam tentang preferensi konsumen di industri pengiriman makanan, memberikan dasar untuk pengembangan strategi pemasaran yang lebih cerdas.

Hasil penelitian ini, jika berhasil, akan memberikan wawasan mendalam tentang preferensi konsumen dalam industri pengiriman makanan di Indonesia. Metode *k-means clustering* diharapkan dapat mengidentifikasi pola yang mungkin terlewatkan, memberikan landasan bagi penyedia layanan untuk merancang strategi pemasaran yang lebih tepat sasaran. Implikasi praktisnya mencakup pengembangan menu yang lebih sesuai dengan preferensi konsumen, meningkatkan daya saing bisnis, dan meningkatkan kepuasan pelanggan. Bagi peneliti, temuan ini dapat memotivasi penelitian lanjutan di bidang analisis data dan perilaku konsumen. Keseluruhannya, penelitian ini diharapkan dapat memberikan kontribusi signifikan pada pemahaman dan pengembangan industri pengiriman makanan di Indonesia.

## 2. TINJAUAN PUSAKA

Menurut [4] membahas mengenai *Customer Data Segmentation on Top 3 Online Food Ordering Applications to Members of Roller Skates Community in Jakarta using K-Means Clustering Method*. Kemajuan dalam teknologi telah memudahkan pengguna untuk dapat melakukan berbagai hal hanya dengan menggunakan perangkat ponsel pintar, salah satunya adalah memesan makanan melalui aplikasi pada ponsel pintar. Tujuan dari penelitian ini adalah untuk mengidentifikasi karakteristik pelanggan, kategori produk, dan kategori pedagang dalam 3 aplikasi pemesanan makanan *online* teratas di Jakarta menggunakan metode *K-Means Clustering*. Terdapat tiga variabel, yaitu karakteristik pelanggan, kategori produk, dan kategori pedagang. Pengolahan data dibantu dengan menggunakan aplikasi SPSS V. 25. Pelanggan setia GoFood termasuk dalam kluster 2. Selain itu, GoFood juga memiliki pelanggan setia dalam kluster 3. Pelanggan setia GrabFood termasuk dalam kluster 4. Selain itu, GrabFood juga memiliki pelanggan setia dalam kluster 5. Pelanggan setia ShopeeFood termasuk dalam kluster 1. Selain itu,

ShopeeFood juga memiliki pelanggan setia dalam kluster 3.

Menurut [5] membahas mengenai *Indonesian Consumers Segmentation in Online Food Purchasing*. Penelitian ini bertujuan untuk menguji segmentasi dan karakteristik konsumen Indonesia selama pandemi Covid-19 dalam pembelian makanan *online*. Dalam penelitian ini, segmentasi psikografis gaya hidup dan perilaku pencarian manfaat dipertimbangkan. Sebuah kuesioner Likert skala lima dibagikan kepada 352 orang Indonesia yang berusia lebih dari 17 tahun dan pernah membeli makanan *online*. Kemudian data dianalisis menggunakan analisis faktor dan analisis kluster. Temuan penelitian membentuk lima faktor gaya hidup: peminat tren, orang praktis, orang yang suka bersosialisasi, kesejahteraan, dan berorientasi pada detail; dan ada tiga faktor pencarian manfaat, yaitu berorientasi pada nilai, orang ekonomis, dan orang yang penuh perhatian. Berdasarkan faktor-faktor yang terbentuk, analisis kluster K-means digunakan, dan menunjukkan bahwa ada tiga kelompok konsumen: rasional, petualang, dan dasar. Menurut [6] membahas mengenai Algoritma *Mean Shift* untuk Menentukan Segmentasi Pelanggan pada Penjualan Toko *Online*. Segmentasi pasar merupakan salah satu hal yang sangat krusial bagi sebuah bisnis ataupun usaha, dengan adanya segmentasi pasar sebuah toko atau perusahaan dapat mengetahui kemampuan daya beli, kebutuhan dan karakteristik pelanggan. Tujuan dari penelitian ini adalah untuk mengetahui nilai segmentasi pelanggan pada penjualan di sebuah toko *online* yang berbasis di Inggris dengan penjualan utamanya merupakan hadiah unik untuk berbagai acara dimana pelanggan toko merupakan pedagang grosir dari berbagai negara. Data mining dengan teknik *clustering* dimanfaatkan dalam penelitian ini. Algoritma yang digunakan untuk membangun kluster adalah algoritma *Mean Shift*, dengan estimasi nilai *bandwith* 1.55, nilai *quantil* = 0.4 dan  $n\_samples = 500$  didapatkan 3 kluster yang divisualisasi menggunakan model scatter plot. Menurut [6] membahas mengenai Penerapan *Data Mining* Pengelompokan Menu Makanan dan Minuman Berdasarkan Tingkat Penjualan Menggunakan Metode K-Means. Dalam usaha meningkatkan penjualan, Dpom *Coffee* menghadapi masalah penumpukan stok bahan yang berlebihan akibat fluktuasi penjualan yang tidak stabil. Kesalahan dalam memprediksi penjualan telah menyebabkan pembelian besar-besaran bahan yang tidak banyak digunakan atau terjual, meninggalkan stok bahan yang melimpah di gudang. Dalam mengatasi permasalahan ini, penelitian ini menerapkan *data mining* dengan fokus pada pengelompokan menu berdasarkan tingkat penjualan menggunakan algoritma K-means. Hasil analisis dataset penjualan di Dpom *Coffee* menghasilkan tiga kluster, masing-masing dengan karakteristik tingkat penjualan rendah, sedang, dan tinggi. Implementasi algoritma K-means ini diharapkan dapat membantu Dpom *Coffee* dalam mengelola stok bahan secara lebih

efektif, mengurangi kasus kekurangan dan kelebihan stok, serta membentuk kebijakan untuk meningkatkan penjualan. Akurasi algoritma K-means dinilai melalui indeks Davies Bouldin, yang mencapai nilai 0.457.

Menurut penelitian [7] membahas mengenai Penerapan *Data Mining* Untuk Menentukan Strategi Penjualan Pada Toko Raja Komputer Menggunakan Metode *Clustering*. Toko Raja Komputer Balikpapan merupakan salah satu toko penjualan laptop yang bergerak dibidang elektronik dengan produk utama seperti laptop, *computer pc*, dan aksesoris lainnya. Dimana perusahaan ini setiap harinya harus memenuhi kebutuhan konsumen dan dituntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi penjualan. Untuk dapat melakukan hal tersebut Toko Raja Komputer membutuhkan strategi – strategi penjualan untuk dapat menarik minat pembeli dan meningkatkan laba atau pendapatan perusahaan. Pada penelitian ini analisa data mining dilakukan dengan teknik *clustering* menggunakan metode K-Means. Dari hasil penelitian yang dilakukan didapatkan hasil dari 1044 dataset daftar penjualan Raja Komputer tahun 2021. Maka didapat kan hasil klusterisasi pada data transaksi penjualan tersebut sebanyak 4 cluster yang mana terdapat 398 data mengikuti cluster 0, 363 data mengikuti cluster 1, 145 data mengikuti cluster 2, dan 138 data mengikuti cluster 3, yang artinya cluster 0 dan 1 didefinisikan sebagai prodak yang harus diperhatikan penyediaan stok dengan jumlah paling banyak sedangkan sedangkan *cluster* 2 dan 3 merupakan produk dengan penyediaan stok di bawah *cluster* 0 dan 1.

Menurut penelitian [9] membahas mengenai Pengaruh Program Promosi dan Ketepatan Waktu Pengiriman Terhadap Keputusan Pembelian Konsumen (Studi Kasus: Layanan Pesan Antar Makanan Aplikasi ShopeeFood di Kota Semarang). Penelitian ini bertujuan menganalisis pengaruh ketepatan waktu pengiriman dan program promosi terhadap keputusan pembelian konsumen Shopee Food di Kota Semarang. Shopee, sebagai *market leader e-commerce* di Indonesia, memiliki aplikasi layanan antar-makanan, ShopeeFood, yang belum mencapai posisi *market leader* di industri tersebut. Program promosi seperti Promo *Payday* dan Promo *Double Date* meningkatkan antusiasme pengguna saat berlangsung, karena potongan harga yang signifikan. Ketepatan waktu pengiriman juga menjadi pertimbangan penting. Penelitian menggunakan metode kuantitatif dengan analisis regresi linear berganda. Hasilnya menunjukkan bahwa ketepatan waktu pengiriman dan program promosi berkontribusi pada keputusan pembelian konsumen. Implikasi praktisnya mendukung ShopeeFood dan layanan serupa untuk meningkatkan kepuasan, membangun loyalitas, dan memperluas pangsa pasar melalui pengoptimalan program promosi dan ketepatan waktu pengiriman.

Penelitian [10] membahas mengenai Pilihan Konsumen Dalam Menggunakan Fitur Go-Food:

Variabel Kegunaan sebagai Moderator. Penelitian ini bertujuan untuk mengevaluasi peran kegunaan dalam memediasi pengaruh pilihan terhadap niat menggunakan fitur Go-Food pada aplikasi Gojek. Dilakukan di Fakultas Ekonomi dan Bisnis Universitas Ngurah Rai, Denpasar, Bali, dengan sampel 100 mahasiswa yang dipilih secara acak. Model persamaan *Struktural Partial Least Square (PLS)* digunakan. Hasil penelitian menunjukkan pilihan berpengaruh positif dan signifikan terhadap kegunaan yang dirasakan. Kegunaan yang dirasakan juga berpengaruh positif dan signifikan terhadap niat menggunakan fitur Go-Food. Variasi makanan juga memiliki pengaruh positif dan signifikan terhadap niat menggunakan fitur tersebut. Kegunaan yang dirasakan berperan sebagai mediator positif dan signifikan dalam hubungan antara pilihan dan niat menggunakan fitur Go-Food pada aplikasi Gojek.

Menurut [11] membahas mengenai *Clustering Minimarket* Untuk Menentukan Jumlah Kebutuhan Pembelian Menggunakan Metode K-Means. Penelitian ini merespon kebutuhan perusahaan terkait pengelolaan data penjualan makanan dan minuman di *Minimarket*. Dengan menggunakan metode *data mining*, penelitian bertujuan mengembangkan sistem clusterisasi untuk menentukan jumlah kebutuhan makanan dan minuman pada setiap *minimarket*. Sistem ini diharapkan dapat meningkatkan efisiensi pengambilan keputusan, memberikan acuan yang efektif, dan memperjelas pola penjualan produk, potensial meningkatkan kinerja bisnis di sektor penjualan makanan dan minuman pada *minimarket*.

Menurut [12] Penerapan Menggunakan Metode Algoritma K-Means Pada Data Stok Toko Baker Old Poris adalah aturan *clustering* menggunakan algoritma K-Means dapat digunakan untuk menentukan kombinasi Toko Baker Old Poris. Pemilik dapat menyediakan produk yang banyak digunakan dengan hasil aturan K-Means Toko Baker Old Poris. Hasil menerapkan metode K-Means untuk menganalisis klasifikasi produk toko baker old poris, didapat Roti kopi isi coklat *topping coffee* dan Roti kopi isi keju *topping coffee* sebagai produk yang sangat laku dan akan menjadi bahan yang sering di stok.

Penelitian membahas [13] mengatakan Perbandingan dilakukan untuk mendapatkan cluster terbaik dengan pemodelan percobaan dari K4 hingga K10. Dimana algoritma *fuzzy c-means* menghasilkan nilai SI 0,2159 dan algoritma *k-medoids* menghasilkan nilai SI 0,2018. Berdasarkan nilai SI tersebut, algoritma *fuzzy c-means* lebih baik dalam penempatan *cluster* dibandingkan dengan algoritma *k-medoids*.

### 3. METODE PENELITIAN

Penelitian analisis data produk GoFood di Indonesia dengan menggunakan metode *k-means clustering* melibatkan serangkaian langkah untuk menggolongkan produk ke dalam kelompok berdasarkan karakteristik tertentu. Pertama, data produk GoFood dikumpulkan, termasuk informasi

seperti nama produk, harga, kategori, *rating*, dan atribut lainnya. Kemudian, dilakukan persiapan data dengan membersihkan data dari nilai yang tidak valid atau duplikat, serta normalisasi jika diperlukan. Selanjutnya, jumlah kluster (K) dipilih, dan titik awal sebagai pusat kluster diinisialisasi secara acak. Proses iteratif dimulai dengan menghitung jarak antara setiap titik data dan pusat kluster, dilanjutkan dengan pengelompokan data berdasarkan pusat kluster terdekat. Pusat kluster diperbarui dengan menghitung rata-rata dari semua titik data dalam kluster, dan iterasi dilakukan hingga konvergensi. Hasil klustering kemudian diinterpretasikan untuk memahami pola atau kelompok produk, serta karakteristik yang mendefinisikan setiap kluster. Validasi hasil dan visualisasi data dilakukan untuk memastikan kebenaran interpretasi. Keseluruhan metodologi ini didokumentasikan secara jelas, memberikan pemahaman mendalam tentang produk GoFood untuk mendukung pengambilan keputusan bisnis dan strategi pemasaran yang lebih efektif.

#### 3.1. Analisis Data

Metode yang digunakan penelitian fokus pada penggunaan metode K-Means untuk menganalisis dan mengelompokkan data terkait data penjualan. Penelitian ini mungkin bertujuan untuk mengidentifikasi atau kelompok dalam data yang dapat memberikan wawasan tentang aspek-aspek transaksi penjualan barang. Dengan menggunakan metode K-Means, penelitian ini akan mencoba memahami lebih dalam struktur data penjualan dan mungkin memberikan pandangan yang dapat mendukung pengambilan keputusan dan tindakan yang lebih baik dalam meningkatkan penjualan produk.

*Cluster* adalah proses pengelompokan objek-objek ke dalam beberapa group yang berbeda, atau mempartisi data set menjadi subset. Formula fungsi objektif dinyatakan dengan persamaan (1):

$$J = \sum_{j=1}^k \sum_{i=1}^n |x_i^{(j)} - C_j|^2 \quad (1)$$

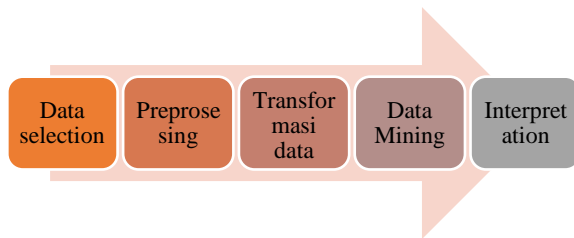
- J : fungsi objektif
- K : jumlah *cluster*
- N : jumlah kasus
- X : case i
- C<sub>j</sub> : *centroid* untuk *cluster j*

*Clustering* berbasis-*centroid* adalah suatu algoritma iteratif di mana gagasan kesamaan diturunkan dengan seberapa dekat suatu titik data dengan *centroid* dari *cluster*.

Nilai *centroid* dapat dicari dengan menggunakan persamaan (2):

$$c = \sum_{i=1}^n X_i/n \quad (2)$$

- C : *centroid* pada *cluster*
- X<sub>i</sub> : titik / objek ke-i
- n : jumlah objek



Gambar 1. Tahapan penelitian

### 3.1.1. Davies Bouldin Index (DBI)

*Davies Bouldin Index (DBI)* diperkenalkan oleh David L. Davies dan Donald W. Bouldin pada tahun 1979. *Davies-Bouldin Index* merupakan salah satu metode yang digunakan untuk mengevaluasi *cluster*. Metode evaluasi *clustering* dalam penelitian ini berdasarkan hasil dari *Davies-Bouldin Index* mengukur seberapa baik kluster terpisah satu sama lain. Semakin rendah nilai *Davies-Bouldin Index*, semakin baik pemisahan antar kluster. Nilai ini dianggap baik jika mendekati 0. Berdasarkan hasil dari penelitian ini bahwa nilai dari *Davies-Bouldin Index* menghasilkan nilai 0.407. Hasil yang diperoleh dalam mencari nilai *Davies-Bouldin Index (DBI)* dinyatakan dari persamaan berikut (3):

$$DBI = \frac{1}{k} \sum_i^k = 1 \max_{i \neq j} (R_{i,j}) \quad (3)$$

$k$  merupakan jumlah *cluster* yang digunakan. Semakin kecil nilai DBI yang diperoleh, maka semakin baik pemisahan antar kluster satu sama lain.

### 3.1.2. Sumber Data

Dalam penelitian ini, sumber data yang digunakan berasal dari data *public* Indonesia food delivery Gofood product list (kaggle.com). Penelitian ini merinci analisis mendalam mengenai tren dan karakteristik pasar layanan pengiriman makanan di Indonesia. Sumber data utama yang digunakan dalam penelitian ini diperoleh dari daftar produk Gofood, sebuah layanan pengiriman makanan yang populer di Indonesia. Informasi yang terkandung dalam dataset ini melibatkan berbagai aspek, termasuk jenis makanan, harga, dan lokasi penyedia layanan. Data ini dikumpulkan secara publik melalui *platform* Kaggle, sebuah sumber terpercaya untuk data ilmiah dan analisis. Penelitian ini bertujuan untuk memberikan wawasan mendalam tentang preferensi konsumen, tren konsumsi, dan dinamika pasar dalam industri pengiriman makanan di Indonesia. Melalui analisis terperinci terhadap dataset yang ada, diharapkan dapat ditemukan pola-pola unik dan informasi berharga yang dapat menjadi dasar untuk pengambilan keputusan yang lebih baik dalam pengembangan strategi bisnis di sektor ini. Dengan memanfaatkan sumber data yang kredibel dan representatif seperti Kaggle, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam pemahaman mendalam tentang pasar layanan pengiriman makanan di Indonesia.

### 3.1.3. Populasi dan Sampel

Populasi dalam penelitian ini merupakan data *Food Delivery GoFood Product List* yang tersedia untuk umum yang dapat ditemukan di Kaggle yang memiliki dataset berjumlah 7770 data. Dataset ini mencakup berbagai jenis data *Food Delivery GoFood Product List*.

Sampel ini merupakan bagian dari populasi yang akan digunakan dalam penelitian untuk deteksi dan dikelompokkan. Metode pengambilan sampel yang digunakan adalah sampel jenuh, yang artinya adalah seluruh populasi dijadikan *sampling* yaitu sebanyak 7770 data.

### 3.2. Teknik Pengumpulan Data

Metode penelitian kepustakaan yang kegiatannya dilakukan adalah mempelajari, mencari dan mengumpulkan data yang berkaitan dengan penelitian tersebut, telah digunakan untuk kegiatan pengumpulan data dalam penelitian ini. Kemudian untuk mengelompokkan daftar produk GoFood yang diperoleh akan diolah berdasarkan metode K-Means yang mengambil semua nilai atribut yang terdapat pada data tersebut.

#### 3.2.1. Data Mining

Proses mencari atau menggali informasi menarik dalam sebuah data dengan jumlah besar yang terpilih dengan menggunakan teknik atau metode tertentu. Penelitian ini menggunakan algoritma *K-Means* untuk mengelompokkan transaksi penjualan produk snack dan minuman berdasarkan *category*. Dalam penelitian ini pengolahan transaksi penjualan menggunakan bantuan *software* RapidMiner dengan menggunakan algoritma *K-Means clustering*. *Data mining* adalah proses analitis yang dirancang untuk mengeksplorasi sejumlah besar data untuk pengetahuan yang berharga, konsisten, dan tersembunyi. [14]

#### 3.2.2. Interpretation/Evaluation

Dalam konteks analisis kluster, evaluasi kualitas kluster sangat penting untuk memahami sejauh mana kluster yang dihasilkan sesuai dengan tujuan analisis dan apakah mereka memberikan wawasan yang berguna. metode evaluasi *clustering* dalam penelitian ini berdasarkan hasil dari *Davies-Bouldin Index* mengukur seberapa baik kluster terpisah satu sama lain. Semakin rendah nilai *Davies-Bouldin*, semakin baik pemisahan antar kluster. Nilai ini dianggap baik jika mendekati 0. Berdasarkan hasil dari penelitian ini bahwa nilai dari *Davies-Bouldin* menghasilkan nilai 0.407.

#### 3.2.3. Knowledge presentation

Tahapan *Knowledge presentation* menampilkan pola informasi yang dihasilkan dari proses algoritma, proses ini memaparkan hasil sehingga menjadi informasi yang mudah dipahami.

#### 4. HASIL DAN PEMBAHASAN

Memuat hasil, pengujian dan pembahasan tentang analisis yang telah dilakukan

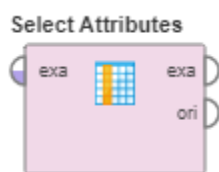
##### 4.1. Data Selection

Pada tahap ini, proses pemilihan data yang relevan dan sesuai dengan tujuan penelitian dilakukan. Data awal yang digunakan dalam penelitian ini yaitu data dataset transaksi penjualan produk *snack* dan minuman. Data awal yang digunakan dalam penelitian ini yaitu data *Food Delivery Gofood Product List*. sebanyak 9722 data yang terdiri dari 9 atribut yaitu *merchan\_name*, *merchant\_area*, *category*, *display*, *product*, *price*, *discount\_price*, *isDiscount*, *description*. Proses seleksi data dilakukan sehingga data hasil proses seleksi data dan pemilihan atribut yang akan dijadikan bahan penelitian ini menjadi 7770 data dan 8 atribut.

Tabel 1. Tabel section atribut

| No | Nama Atribut   | Type        |
|----|----------------|-------------|
| 1  | merchant_name  | Polynominal |
| 2  | merchant_area  | Polynominal |
| 3  | Category       | Polynominal |
| 4  | Display        | Polynominal |
| 5  | Product        | Polynominal |
| 6  | Price          | Polynominal |
| 7  | discount_price | Integer     |
| 8  | isDiscountt    | Integer     |

Penggunaan *select atribut* seperti pada proses yang ada pada gambar 2 untuk menyeleksi atribut-atribut yang akan digunakan dalam penelitian ini, dalam penelitian ini data yang akan digunakan setelah dilakukan proses seleksi menjadi 8 atribut yang dijadikan bahan penelitian. Kemudian dilanjutkan dengan penerapan operator *replace missing value*. Adapun model prosesnya seperti pada gambar 3



Gambar 2. Select atribut

##### 4.2. Preprocessing

*Preprocessing*, atau pra-pemrosesan data, adalah tahapan yang dilakukan dalam analisis data menggunakan metode *k-means clustering*. Langkah-langkah ini membantu mempersiapkan data agar sesuai dengan kebutuhan analisis. Beberapa langkah *preprocessing* yang menggunakan operator *replace missing value*:



Gambar 3. Replace missing value

##### 4.3. Pembersihan Data

Identifikasi dan tangani data yang hilang atau tidak lengkap, serta deteksi dan penanganan *outlier* jika diperlukan. Data yang bersih akan menghasilkan hasil *clustering* yang lebih akurat. Pada tahapan ini pembersihan data dilakukan dengan cara menggunakan operator *replace missing value*, sebelum dilakukan pembersihan data terdapat data yang *missing* pada beberapa atribut dan data statistiknya dapat dilihat pada gambar 4 kemudian hasilnya setelah dilakukan proses pembersihan data dapat dilihat pada gambar 5

Gambar 4. Statistik datasetgofood

Dilihat dari hasil data statistik datasetgofood terlihat bahwa terdapat data yang *missing* pada atribut *price* sebanyak 47 data.

Gambar 5. Statistik datasetgofood

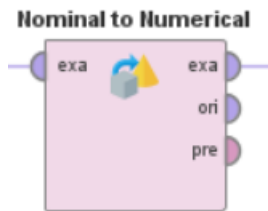
Setelah diketahui banyak data yang *missing* maka dalam penelitian ini dilakukan proses *replace missing value* dan setelah dilakukan hasilnya seperti terlihat pada gambar 4 bahwa dari keseluruhan atribut sudah tidak ada data yang *missing*.

Berdasarkan gambar 4 model proses *replace missing value* digunakan untuk menangani data yang *missing* pada atribut *price*.



#### 4.4. Transformasi Data

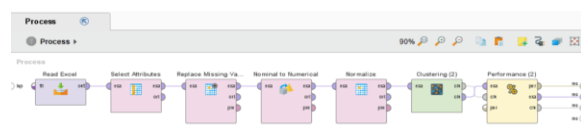
Proses tranformasi data kedalam bentuk format tertentu sehingga data tersebut sesuai untuk proses *data mining*. Dalam konteks penggunaan *k-means* pada RapidMiner untuk data yang memiliki variabel kategori (nominal), transformasi nominal ke numerik dapat menjadi langkah yang diperlukan menggunakan *k-means* di RapidMiner. *K-means* bekerja dengan baik pada data numerik. Jika dataset memiliki variabel kategori, perlu mengonversinya menjadi bentuk numerik agar dapat diolah oleh algoritma *k-means*. Operator *nominal to numeric* yang digunakan dalam penelitian ini seperti pada gambar 6.



Gambar 6. Nominal to numeric

#### 4.5. Model Data Mining

Proses mencari atau menggali informasi menarik dalam sebuah data dengan jumlah besar yang terpilih dengan menggunakan teknik atau metode tertentu. Penelitian ini menggunakan algoritma *K-Means* untuk mengelompokkan *datasetgofood* berdasarkan *category*. Dalam penelitian ini pengolahan *datasetgofood* menggunakan bantuan *software* RapidMiner dengan menggunakan algoritma *K-Means clustering*. Adapun tahapan *data mining* dalam penelitian ini adalah dimulai dengan mengaktifkan tools *rapidminer* dan setelah dijalankan maka akan tampil seperti pada gambar 7.



Gambar 7. Model proses *k-means*

Proses mencari atau menggali informasi menarik dalam sebuah data dengan jumlah besar yang terpilih dengan menggunakan teknik atau metode tertentu. Penelitian ini menggunakan algoritma *K-Means* untuk mengelompokkan *datasetgofood* berdasarkan *category*. Dalam penelitian ini pengolahan *datasetgofood* menggunakan bantuan *software* RapidMiner dengan menggunakan algoritma *K-Means clustering*. Adapun tahapan *datamining* dalam penelitian ini adalah dimulai dengan mengaktifkan *tools* RapidMiner dan setelah dijalankan maka akan tampil seperti pada gambar 7.

Dalam konteks pemodelan mesin, seperti pada gambar 7 merupakan model proses algoritma *K-Means* yang digunakan untuk melakukan klastering terhadap *datasetgofood* yang terkait dengan penelitian ini. Hasil dari model proses tersebut peneliti telah melakukan uji

*k=2* sampai dengan *k=10* dan hasilnya dapat dilihat pada tabel 2.

Tabel 2. Hasil *davies bouldin indeks (DBi)*

| K  | DBi           | Cluster                                                                                                                                                                                                                      |
|----|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2  | -0.927        | Cluster 0: 7768 items<br>Cluster 1: 2 items                                                                                                                                                                                  |
| 3  | -0.664        | Cluster 0: 7766 items<br>Cluster 1: 3 items<br>Cluster 2: 1 items                                                                                                                                                            |
| 4  | -1.100        | Cluster 0: 7758 items<br>Cluster 1: 7 items<br>Cluster 2: 4 items<br>Cluster 3: 1 items                                                                                                                                      |
| 5  | -2.053        | Cluster 0: 7732 items<br>Cluster 1: 1 items<br>Cluster 2: 1 items<br>Cluster 3: 33 items<br>Cluster 4: 3 items                                                                                                               |
| 6  | -1.440        | Cluster 0: 1 items<br>Cluster 1: 7736 items<br>Cluster 2: 23 items<br>Cluster 3: 8 items<br>Cluster 4: 1 items<br>Cluster 5: 1 items                                                                                         |
| 7  | -0.818        | Cluster 0: 7751 items<br>Cluster 1: 2 items<br>Cluster 2: 1 items<br>Cluster 3: 8 items<br>Cluster 4: 1 items<br>Cluster 5: 6 items<br>Cluster 6: 1 items                                                                    |
| 8  | <b>-0.407</b> | <b>Cluster 0: 7763 items<br/>Cluster 1: 1 items<br/>Cluster 2: 1 items<br/>Cluster 3: 1 items<br/>Cluster 4: 1 items<br/>Cluster 5: 1 items<br/>Cluster 6: 1 items<br/>Cluster 7: 1 items</b>                                |
| 9  | -1.501        | Cluster 0: 7670 items<br>Cluster 1: 1 items<br>Cluster 2: 1 items<br>Cluster 3: 1 items<br>Cluster 4: 1 items<br>Cluster 5: 93 items<br>Cluster 6: 1 items<br>Cluster 7: 1 items<br>Cluster 8: 1 items                       |
| 10 | -1.127        | Cluster 0: 59 items<br>Cluster 1: 1 items<br>Cluster 2: 1 items<br>Cluster 3: 1 items<br>Cluster 4: 7703 items<br>Cluster 5: 1 items<br>Cluster 6: 1 items<br>Cluster 7: 1 items<br>Cluster 8: 1 items<br>Cluster 9: 1 items |

Berdasarkan hasil dari tabel 2 hasil *Davies Bouldin Indeks* merupakan kluster yang optimal pada *k=8* dalam penelitian ini dengan menghasilkan nilai *DBi* = -0.407 dan untuk melihat hasil uji *k=2* sampai dengan *k=10*.

#### 4.6. Evaluasi Model

Dalam konteks analisis kluster, evaluasi kualitas kluster sangat penting untuk memahami sejauh mana kluster yang dihasilkan sesuai dengan tujuan analisis dan apakah mereka memberikan wawasan yang berguna. Metode evaluasi *clustering* dalam penelitian ini berdasarkan hasil dari *Davies-Bouldin Index* mengukur seberapa baik kluster terpisah satu sama lain. Semakin rendah nilai *Davies-Bouldin*, semakin baik pemisahan antar kluster. Nilai ini dianggap baik jika mendekati 0. Berdasarkan hasil dari penelitian ini bahwa nilai dari *Davies-Bouldin* menghasilkan nilai 0.407.

#### 4.7. Pembahasan

Hasil dari penelitian yang dilakukan terhadap data *food delivery gofood product list* menggunakan metode algoritma k-means hasil kluster terbaik terdapat pada k8 dengan nilai *Davies-Bouldin Index* (DBI) sebesar -0.407 dan setiap *cluster* memiliki jumlah item yang berbeda. *Cluster* 0 merupakan *cluster* terbesar dan mungkin mencakup kategori atau jenis produk makanan yang paling umum atau luas dengan memiliki 7763 item, sementara *cluster* 1 hingga *cluster* 7 masing-masing memiliki 1 item.

#### 4.8. Bagaimana mengimplementasikan pengelompokan data *Food Delivery Gofood Product List* menggunakan metode *K-Means Clustering*?

Implementasi pengelompokan data *Food Delivery GoFood Product List* menggunakan metode *K-Means Clustering* melibatkan langkah-langkah yang terstruktur. Pertama, data produk GoFood dipersiapkan dengan membersihkan data dan memilih fitur yang relevan seperti harga, jenis makanan, *rating*, dan waktu pengiriman. Selanjutnya, jumlah *cluster* (*K*) ditentukan, seringkali menggunakan metode seperti *Elbow Method* atau *Silhouette Score*. Kemudian, algoritma *K-Means* diterapkan dengan inisialisasi *centroid* secara acak, diikuti oleh iterasi penugasan dan pembaruan *centroid* hingga konvergensi. Hasil pengelompokan dievaluasi dengan visualisasi menggunakan *scatter plot* untuk memperlihatkan pola *cluster*, dan statistik deskriptif untuk menganalisis karakteristik setiap *cluster* seperti rata-rata harga atau *rating*. Evaluasi lebih lanjut dilakukan dengan mengukur kesamaan *intra-cluster* dan pemisahan antar-*cluster*. Interpretasi dilakukan dengan menganalisis *centroid-cluster* untuk mendapatkan gambaran fitur-fitur yang paling membedakan antar *cluster*. Validasi menggunakan metrik seperti *Silhouette Score* membantu mengukur kualitas pengelompokan. Hasil analisis memberikan wawasan tentang segmentasi produk, strategi pemasaran, dan pengembangan produk yang sesuai dengan preferensi konsumen dalam setiap *cluster*. Perbaikan model dapat dilakukan jika hasil analisis menunjukkan *cluster* yang tidak sesuai, sehingga implementasi *K-Means Clustering* pada data *GoFood Product List* menjadi

langkah penting dalam meningkatkan efisiensi dan kepuasan pengguna layanan pengiriman makanan. Hasil implementasi *K-Means Clustering* dari penelitian yang dilakukan terhadap data *food delivery gofood product list* menggunakan metode algoritma k-means diperoleh terbaik sebanyak 8 *cluster* dengan nilai *Davies-Bouldin Index* (DBI) =-0.407, yaitu dengan anggota *cluster* 0=7763 items, *cluster* 1=1 items, *cluster* 2=1 items, *cluster* 3=1 items, *cluster* 4=1 items, *cluster* 5=1 items, *cluster* 6=1 items dan *cluster* 7=1 items. Setiap *cluster* memiliki jumlah item yang berbeda. *Cluster* 0 merupakan *cluster* terbesar dan mungkin mencakup kategori atau jenis produk makanan yang paling umum atau luas dengan memiliki 7763 item, sementara *cluster* 1 hingga *cluster* 7 masing-masing memiliki 1 item.

#### 4.9. Bagaimana cara menganalisa hasil implementasi pengelompokan data *Food Delivery Gofood Product List* menggunakan metode *K-Means Clustering*?

Analisis hasil implementasi pengelompokan data *Food Delivery GoFood Product List* dengan metode *K-Means Clustering* dilakukan melalui beberapa langkah yang sistematis. Pertama-tama, hasil *clustering* dievaluasi dengan menggunakan visualisasi, seperti *scatter plot* untuk memperlihatkan pola-pola *cluster* dalam ruang fitur. Hal ini membantu dalam mengidentifikasi seberapa baik data produk terkelompok dalam *cluster-cluster* yang berbeda. Selanjutnya, karakteristik masing-masing *cluster* dianalisis dengan menggunakan statistik deskriptif, seperti rata-rata harga, *rating*, atau waktu pengiriman. Perbandingan antara *cluster-cluster* juga dapat dilakukan dengan mengamati *centroid-cluster*, yaitu nilai rata-rata dari setiap fitur pada *cluster*. Hal ini memberikan gambaran tentang seberapa homogen setiap *cluster* dan seberapa jauh pemisahan antar *cluster*. Selain itu, analisis secara visual dengan membandingkan karakteristik setiap *cluster*, seperti melalui grafik atau tabel, dapat memberikan wawasan yang lebih mendalam tentang perbedaan antara kelompok produk. Hasil analisis ini kemudian dapat diinterpretasikan untuk mengidentifikasi pola-pola penting, segmentasi pasar yang relevan, dan strategi bisnis yang dapat diimplementasikan berdasarkan setiap *cluster*. Dengan demikian, analisis hasil implementasi *K-Means Clustering* pada data *Food Delivery GoFood Product List* memberikan pemahaman yang komprehensif tentang karakteristik produk, segmentasi pasar yang mungkin, dan langkah-langkah yang dapat diambil untuk meningkatkan efisiensi dan keberhasilan layanan pengiriman makanan.

Dalam hasil pengelompokan menggunakan metode *K-Means*, terdapat 8 *cluster* dengan nilai *Davies-Bouldin Index* (DBI) sebesar -0.407. Berikut adalah interpretasi dari hasil tersebut: Jumlah *Cluster* (*k*): Model *K-Means* menghasilkan 8 *cluster* berdasarkan data produk makanan. Nilai DBI yang



negatif (-0.407) dapat dianggap baik. DBI adalah ukuran yang mengukur seberapa baik *cluster-cluster* tersebut dipisahkan satu sama lain. Semakin rendah nilai DBI, semakin baik pembagian *cluster*-nya. Dalam hal ini, nilai negatif menunjukkan bahwa *cluster-cluster* tersebut relatif baik dipisahkan. Ukuran *Cluster*: Setiap *cluster* memiliki jumlah item yang berbeda. *Cluster* 0 memiliki 7763 item, sementara *cluster* 1 hingga *cluster* 7 masing-masing memiliki 1 item. *Cluster* 0 adalah *cluster* terbesar dan mungkin mencakup kategori atau jenis produk makanan yang paling umum atau luas. Interpretasi *Cluster*: Perlu lakukan analisis lebih lanjut untuk memahami karakteristik masing-masing *cluster*. *Cluster* yang memiliki jumlah item hanya 1 mungkin perlu diperhatikan lebih cermat untuk melihat apakah ada kesamaan atau perbedaan yang signifikan dalam atribut produk makanan di dalamnya.

## 5. KESIMPULAN DAN SARAN

Berdasarkan pengelompokan data *Food Delivery GoFood Product List* menggunakan metode *K-Means Clustering*, hasil *Davies Bouldin Indeks (DBI)* merupakan kluster yang optimal pada  $k=8$  dalam penelitian ini dengan menghasilkan nilai  $DBI = -0.407$ . Setiap *cluster* memiliki jumlah item yang berbeda. *Cluster* 0 memiliki 7763 item, sementara *cluster* 1 hingga *cluster* 7 masing-masing memiliki 1 item. *Cluster* 0 adalah *cluster* terbesar dan mungkin mencakup kategori atau jenis produk makanan yang paling umum atau luas.

Penelitian ini secara efektif dapat mengeksplorasi penggunaan metode *k-means clustering* dalam mengelompokkan data *Food Delivery* dari *GoFood Product List*. Hasil *clustering* berhasil mengidentifikasi pola konsumen serupa dalam kelompok produk makanan, memberikan pemahaman yang lebih dalam tentang preferensi konsumen di sektor pengiriman makanan. Selanjutnya, hasil penelitian ini dapat menjadi alat strategis untuk meningkatkan personalisasi layanan, merancang strategi pemasaran yang lebih efektif, dan menghasilkan inovasi produk yang lebih sesuai dengan preferensi pasar.

Saran untuk peningkatan layanan *Food Delivery GoFood Product List* yaitu perlu dikembangkan sistem rekomendasi yang lebih canggih berdasarkan hasil *clustering* untuk menyajikan pilihan produk yang lebih spesifik sesuai dengan preferensi konsumen. Kerjasama yang erat dengan restoran juga perlu ditingkatkan untuk mengadaptasi penawaran produk mereka sesuai dengan karakteristik kelompok konsumen tertentu. Dan pelatihan karyawan terkait implementasi hasil *clustering* juga menjadi penting agar pemahaman yang mendalam tentang preferensi konsumen dapat diimplementasikan secara efektif

dalam layanan sehari-hari. Melalui kerjasama industri yang lebih luas, termasuk dengan produsen makanan atau penyedia layanan logistik, GoFood dapat memperluas dan meningkatkan variasi produk yang ditawarkan.

## DAFTAR PUSTAKA

- [1] Oktaviano. A, "Pengaruh Program Promosi dan Ketepatan Waktu Pengiriman Terhadap Keputusan Pembelian Konsumen (Studi Kasus: Layanan Pesan Antar Makanan Aplikasi ShopeeFood di Kota Semarang)," 2023.
- [2] N. Arum, "Bisnis Online: Preferensi Konsumen Terhadap Layanan Online Food Delivery," 2020.
- [3] A. Agustriyani, "Faktor yang Mempengaruhi Minat Pembelian Konsumen Pada Aplikasi Food Delivery Selama Pandemi COVID-19," 2021.
- [4] A. D. Resti, "Segmentasi Data Pelanggan pada Top 3 Aplikasi Pesan Makanan Online pada Anggota Komunitas Sepatu Roda di Jakarta Menggunakan Metode K-Means Clustering," 2023
- [5] A. S. B. Santoso, "Indonesian Consumers Segmentation in Online Food Purchasing," 2023.
- [6] R. Reliovani, "Algoritma Mean Shift untuk Menentukan Segmentasi Pelanggan pada Penjualan Toko Online," 2021
- [7] G. Triyandana, "Penerapan Data Mining Pengelompokan Menu Makanan dan Minuman Berdasarkan Tingkat Penjualan Menggunakan Metode K-Means," 2022
- [8] A. S. Wahyuni, "Penerapan Data Mining Untuk Menentukan Strategi Penjualan Pada Toko Raja Komputer Menggunakan Metode Clustering," 2022.
- [9] Oktaviano. A, "Pengaruh Program Promosi dan Ketepatan Waktu Pengiriman Terhadap Keputusan Pembelian Konsumen (Studi Kasus: Layanan Pesan Antar Makanan Aplikasi ShopeeFood di Kota Semarang)," 2023.
- [10] L. Erynayati, "Pilihan Konsumen Dalam Menggunakan Fitur Go-Food: Variabel Kegunaan sebagai Moderator," 2021
- [11] F. K. Alvisan, "Clustering Minimarket Untuk Menentukan Jumlah Kebutuhan Pembelian Menggunakan Metode K-Means," 2021
- [12] F. F. Ramadhan, "Implementasi Algoritma Metode K-Means untuk Analisis Stok Barang pada Baker Old Poris," 2022.
- [13] A. Jaini, "Perbandingan Algoritma Fuzzy C-Means dan K-Medoids untuk Pengelompokan Data Penjualan pada 212 Mart," 2020
- [14] B. I. Nugroho, "Tinjauan Pustaka Sistematis: Penerapan Data Mining Metode Klasifikasi Untuk Menganalisa Penyalahgunaan Sosial Media," 2022.