

BAB 2

TINJAUAN PUSTAKA

2.1. DATA MINING

Mencari informasi atau pola menarik dalam data tertentu dengan menggunakan suatu pemodelan, metode, teknik, atau algoritma tertentu adalah tahapan paling penting dalam proses penemuan pengetahuan (KDD). Pemilihan metode, teknik, atau algoritma maupun pemodelan yang tepat sangat bergantung pada proses dan tujuan penemuan pengetahuan dalam data (Vithon Carelsa dkk., 2023). Salah satu teknik analisa Data *mining* adalah analisis kelompok (*cluster analysis*) yang lebih dikenal dengan *Clustering*. *Clustering* merupakan metode analisis data yang tujuannya mengelompokkan data dengan karakteristik yang sama ke suatu wilayah yang sama (Sahara Lubis dkk., 2023). Data *mining* ini dapat dipilih sebagai sistem yang tepat untuk kemajuan perusahaan yang dapat menemukan informasi yang berada di gudang data perusahaan yang terletak pada penyimpanan data komputer dan data bisnis tersebut dapat bermanfaat sebagai informasi dalam setiap langkah pembangunan (Astuti, 2023).

Berikut ini merupakan proses KDD pada data mining secara garis besar :

1. Data Selection merupakan tahapan pemilihan (seleksi) data. Tahapan ini perlu dilakukan sebelum penggalian informasi di lakukan. Pada tahap ini memilih kumpulan data yang akan digunakan untuk menemukan pengetahuan.
2. Preprocessing/cleaning, dilakukan pembersihan data yang mencakup pembuangan data yang duplikat, noise, kesalahan cetak/tipografi serta menangani data yang tidak konsisten dan relevan.
3. Transformation, data disesuaikan dengan format ekstensi yang sesuai untuk proses data mining. Ini dilakukan karena beberapa metode data *mining* memerlukan format tertentu sebelum diproses..
4. Data mining, menemukan pola atau informasi menarik dalam data melalui penggunaan teknik tertentu. Data mining menggunakan banyak teknik atau algoritma bergantung pada tujuan dan proses KDD secara keseluruhan, pendekatan atau *Algoritme* yang sesuai akan dipilih.
5. Interpretation/evaluation, tahap ini termasuk dalam proses KDD yang dikenal sebagai interpretasi. Pola yang dihasilkan dari proses data *mining* harus ditampilkan sehingga pihak yang berkepentingan dapat memahaminya. Pada titik ini juga dilakukan pemeriksaan ulang untuk memastikan apakah pola atau informasi yang

dikumpulkan bertentangan dengan fakta atau hipotesis sebelumnya (Herlinda Mawarni, 2022).

2.2. PELANGGAN

Pelanggan atau customer merupakan salah satu aset yang dimiliki oleh perusahaan. Perusahaan harus menjaga pelanggan yang dimiliki untuk dapat terus berjalan dan berkembang. Semakin banyak dan loyal pelanggan yang dimiliki oleh perusahaan, maka semakin baik untuk perusahaan (Saputra & Arianty, 2019). Namun sifat pelanggan yang selalu pilih-pilih membuat perusahaan kesulitan dalam membedakan pelanggan yang dapat memberikan keuntungan besar bagi perusahaan dan yang tidak atau kurang menguntungkan bagi Perusahaan (Pramudiansyah, 2021). Upaya untuk mengetahui loyalitas pelanggan bagi perusahaan yaitu dengan konsep segmentasi pelanggan (Pailan et al., 2022).

Segmentasi pelanggan adalah proses membagi pelanggan menjadi sub kelompok yang berbeda, bermakna, dan homogen berdasarkan atribut dan karakteristiknya (Harahap et al., 2022). Hal tersebut memungkinkan suatu perusahaan dapat memahami atau mempelajari pelanggan mereka dan membangun strategi untuk menjalin hubungan dengan pelanggan sesuai dengan karakteristik pelanggan tersebut (Suharti et al., 2022).

2.3. RFM

Model RFM adalah ekspresi numerik perilaku pelanggan dan sangat efektif untuk menentukan pelanggan utama. Namun, definisi dan perhitungan variabel-variabel ini dapat berubah tergantung pada masalahnya (Wicaksono, 2020). Analisis RFM merupakan suatu Model perhitungan yang terdiri dari tiga Atribut domain yang memperhatikan transaksi pelanggan berdasarkan transaksi terakhir/keterkinian (Recency), jumlah transaksi (Frequency), serta nominal dari transaksi (Monetary) (Nuraeni et al., 2023). Tujuan dari RFM adalah untuk meramalkan perilaku Pelanggan di masa depan agar dapat mengarahkan keputusan segmentasi yang lebih baik. Model RFM ini merupakan metode yang sudah lama dan populer untuk mengukur hubungan dengan pelanggan. Dimensi model RFM menurut (Pramudiansyah, 2021) meliputi :

1. Recency adalah mengukur nilai pelanggan dengan melihat perilaku Pelanggan yang berkenaan dengan pembelian yang dilakukannya paling akhir. Informasi terpenting yang tidak boleh dilewatkan berkenaan dengan resensi adalah tanggal pembelian terakhir, karena proses ini bergantung pada tanggal

pembelian terakhir, semakin dekat proses pembelian terakhir maka pelanggan tersebut semakin loyal.

2. Frequency adalah mengukur nilai pelanggan dengan melihat perilaku Pelanggan yang berkenaan dengan aktifitas transaksi yang dilakukan oleh Pelanggan dalam satu periode.
3. Monetary adalah mengukur nilai pelanggan dengan melihat perilaku Pelanggan yang berkenaan dengan nilai pembelian yang dilakukan oleh pelanggan dalam periode tertentu. Semakin tinggi nilai pembelian maka semakin tinggi tingkat loyalitas pelanggan.

Setelah didapatkan nilai dari masing-masing atribut Recency, Frequency dan monetary selanjutnya yang perlu dilakukan adalah memberikan score kuantitatif untuk masing-masing nilai tersebut agar bisa dilakukan pemrosesan di tahap selanjutnya. Pemberian nilai kuantitatif ini digunakan dengan menggunakan rentang skala 1-5 yang memiliki rincian seperti yang ditentukan pada Tabel tersebut (Pailan et al., 2022).

2.4. NORMALISASI MIN-MAX

Metode *Min-Max* merupakan metode normalisasi sederhana dengan melakukan transformasi terhadap data asli. *Min-Mix* akan menyesuaikan batas yang ditentukan dengan menghubungkan pada data asli. Teknik normalisasi ini melakukan transformasi sebuah atribut numerik dalam range atau skala yang lebih kecil seperti 0.0 sampai 1.0 (Nugroho & Adhinata, 2022). Angka tersebut menunjukkan bahwa batas terendah 0.0 sedangkan batas tertinggi 1.0. Menurut (Pramudiansyah, 2021) Normalisasi Min-Max dapat dihitung dengan persamaan (2.1) :

$$v^1 = \frac{v - \min_a}{\max_a - \min_a} (\text{newmax}_a - \text{newmin}_a) + \text{newmin}_a \quad (2.1)$$

Keterangan :

v_1	= nilai setelah dinormalisasi.
v	= nilai sebelum dinormalisasi.
$\min_a$	= nilai minimal setiap variable.
$\max_a$	= nilai maksimal setiap variable.
newmax_a	= rentan maksimal x dengan nilai 1.
newmin_a	= rentan nilai x dengan nilai 0.

2.5. EUCLIDEAN DISTANCE

Euclidian Distance merupakan salah satu metode perhitungan jarak yang digunakan untuk mengukur jarak dari 2 (dua) buah titik yang dalam euclidian space (meliputi bidang euclidian dua dimensi, tiga dimensi atau bahkan lebih)(Mohammad Ferdiansyah, 2024) . Rumus *Pythagoras* dapat digunakan untuk menghitung jarak *Euclidean*, yaitu jarak antara dua titik. Perhitungan menggunakan geometri *Euclidean* memberikan hasil yang positif. Dengan memilih nilai terendah, kesamaan yang diukur menggunakan jarak *Euclidean* ditentukan (Rizaldi et al., 2022). Jika hasil perbandingan dan perhitungan dengan jarak *Euclid* adalah nilai terendah atau bahkan sangat mendekati 0, maka dua item dianggap sebanding (Herlinda Mawarni, 2022).

Berikut adalah persamaan *Euclidean Distance* pada Persamaan (2.1)

$$d(x_i, y_i) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.2)$$

Keterangan :

$D(x_i, y_i)$ = jarak antara clustering x dengan titik centroid y pada data ke- i .

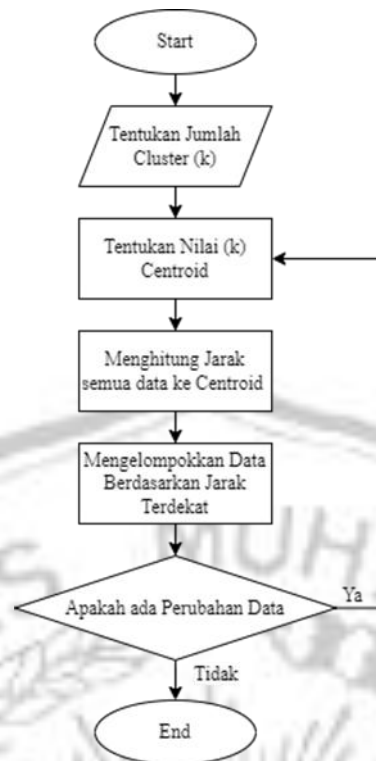
x_i = bobot ke- i pada cluster yang ingin dicari jaraknya.

y_i = bobot data ke- i pada titik centroid.

n = merupakan jumlah data.

2.6. ALGORITMA K-MEANS

K-Means yaitu salah satu dari metode pengelompokan data nonhierarki (sekatan) yang dapat mempartisi data kedalam bentuk dua kelompok ataupun lebih. Metoda tersebut akan mempartisi data kedalam suatu kelompok dimana data yang berkarakteristik sama akan dimasukkan kedalam satu kelompok sama sedangkan data yang memiliki karakteristik yang berbeda akan dikelompokkan kedalam kelompok lainnya. Tujuan dari pengelompokan yaitu untuk meminimalkan dari fungsi objektif yang diset dalam proses pengelompokan, pada umumnya akan berusaha meminimalkan variasi didalam suatu kelompok dan memaksimalkan variasi antar kelompok (Manalu & Gunadi, 2022).



Gambar 2.1 Flowchart K-Means Clustering

Gambar 2.1 menjelaskan tahapan *algoritme K-Means* dimana tahap pertama adalah dengan menentukan banyaknya *cluster* (k) yang akan digunakan. Setelah itu menentukan *centroid* awal secara acak yang diambil dari dataset. Langkah selanjutnya adalah menghitung jarak setiap data ke masing-masing titik *centroid* dengan menggunakan *Euclidean Distance*. Langkah selanjutnya adalah mengelompokkan data tersebut berdasarkan jarak terdekat antara data dengan *centroid*. Kemudian menentukan nilai *centroid* baru dengan menghitung rata-rata dari *cluster*. Setelah itu kita ulangi tahapan menghitung jarak menggunakan teori *Euclidean Distance* hingga tahapan menentukan nilai *cluster* baru. Tahap pengulangan ini akan dihentikan apabila sudah tidak ada perubahan data dalam suatu *cluster*. Untuk menentukan nilai *centroid* baru dapat dilakukan dengan menggunakan persamaan (2.2)

$$C_k = \frac{1}{n_k} \sum_{i=1}^n d_i \quad (2.3)$$

Keterangan :

n_k = jumlah semua data yang ada didalam *cluster* (k).

d_i = jumlah dari setiap *cluster*

2.7. *SILHOUTTE COEFFICIENT*

Silhouette merupakan salah satu metode yang digunakan untuk menguji kualitas klaster yang dihasilkan dari proses *clustering*. Metode ini menghitung rata-rata nilai setiap titik pada himpunan data. Perhitungan nilai setiap titik adalah selisih nilai separation dan compactness yang dibagi dengan maksimum antara keduanya (Satriawan et al., n.d.). Adapun langkah-langkah untuk menghitung nilai *Silhouette Coefficient* sebagai berikut.

Menghitung rata-rata jarak suatu data dengan data lain dalam suatu cluster yang sama menggunakan persamaan (2.3)

$$a(i) = \frac{1}{|A| - 1} \sum_{j \in A, j \neq i} d(i, j) \quad (2.4)$$

Keterangan :

$a(i)$ = Perbedaan rata-rata pada data (i) ke semua data lain di *cluster A*

$d(i, j)$ = Jarak antara data i dan data j

Menghitung rata-rata jarak data tersebut dengan semua data di *cluster* lain menggunakan persamaan (2.4)

$$d(i, C) = \frac{1}{|C|} \sum_{j \in C} d(i, j) \quad (2.5)$$

Keterangan :

$d(i, C)$ = Perbedaan rata-rata pada data (i) ke seluruh data C

C = Cluster lain selain A

Memilih nilai jarak yang paling minimum menggunakan persamaan (2.5)

$$b(i) = \min_{c \neq A} d(i, j) \quad (2.6)$$

Keterangan :

$b(i)$ = Nilai minimum jarak rata-rata data ke- i dengan semua data di *cluster* berbeda.

Menghitung nilai *Silhouette Coefficient* dengan menggunakan persamaan (2.6)

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2.7)$$

Keterangan :

$a(i)$ = Perbedaan rata-rata pada data (i) ke semua data lain di *cluster A*

$b(i)$ = Perbedaan rata-rata pada data (i) ke semua data pada *cluster* berbeda

$s(i)$ = Nilai *Silhouette Coefficient*

Range nilai $s(i)$ yakni antara -1 dan 1, interpretasi nilai tersebut yaitu:

$s(i)$ = -1: data ke- i digolongkan lemah (dekat pada *cluster B* daripada *A*)

$s(i)$ = 0: data ke- i berada di tengah dua kluster (*A* dan *B*)

$s(i)$ = 1: data ke- i digolongkan baik

Intepretasi nilai *Silhouette Coefficient* juga ditunjukkan dalam Tabel 2.1

Tabel 2.1 Interpretasi nilai *Silhouette Coefficient*

<i>Silhouette Coefficient</i>	Interpretasi
≤ 0.25	Tidak terstruktur
0.26-0.50	Hasil struktur lemah
0.51-0.70	Hasil struktur baik
0.71-1.00	Hasil struktur kuat

2.8. PENELITIAN TERKAIT

Tabel 2.2 Penelitian Terkait

Peneliti	Judul	Metode	Objek Penelitian	Hasil
Widiyanto, A T 2021	KONSTELASI: Konvergensi Teknologi dan Sistem Informasi Segmentasi Pelanggan Berdasarkan	metode <i>clustering K-Means</i>	797 data transaksi clean	sistem segmentasi pelanggan dari 29 kali pengujian, terdapat 29 pelanggan (100%) dengan karakteristik yang dihasilkan sistem sesuai dengan pengetahuan user,

	Analisis RFM Menggunakan Algoritma K-Means Sebagai Dasar Strategi Pemasaran (Studi Kasus PT Coversuper Indonesia Global)			Opelanggan (0%) tidak sesuai dengan pengetahuan user.
Rumiarti 2020	Segmentasi Pelanggan Pada Customer Relationship Management Di Perusahaan Ritel: Studi Kasus Pt Gramedia Asri Media	<i>Algoritma K-MEANS</i>	banyaknya jenis buku yang di-beli	Jumlah cluster optimal pada segmentasi nilai pelanggan yang dihasilkan dalam penelitian ini berdasarkan pada nilai tertinggi silhouette method dan calinski-harabasz index. Nilai average silhou-ette pada dua cluster adalah 0.42. Nilai Calinski Harabasz index pada dua cluster adalah 1.31.
Resty Awaliah Febrianty 2021	Segmentasi Penjualan Obat Di Apotek Menggunakan Metode K-Means	Metode K-means	Data Penjualan Obat di Apotek Anugrah Farma Padalarang	Peneraapan metode k-means dapat diterapkan pada data penjualan di apotek Anugrah Farma Padalarang dengan tujuan mengelompokkan data

				penjualan obat tersebut ke dalam empat kelompok, yaitu kelompok penjualan sangat tinggi, penjualan tinggi, penjualan sedang dan penjualan rendah.
Audila Gumanty Widinastia 2021	Sistem Segmentasi Loyalitas Pelanggan Berbasis Rfm Menggunakan Pam Di Laundry Forten's	Metode <i>Partitioning Around Medoids</i>	Data Transaksi Januari Tahun 2018 Sampai Desember 2018 Dari Data Pelanggan Yang Berjumlah 7537 Kemudian Dijadikan Data Transaksi Pelanggan Yang Berjumlah 1089 Data.	Iterasi Pertama Dengan Total Jarak Terkecil Keseluruhan Data = 251,40. Pada Iterasi Kedua Dengan Total Jarak Terkecil Keseluruhan Data = 246,35 Serta Pada Iterasi Ke Tiga Total Jarak Terkecil Keseluruhan Data = 265,87 Sehingga Dapat Disimpulkan Bahwa Hasil Total Jarak Keseluruhan Data Lebih Besar Dari Total Jarak Sebelumnya Maka Proses Iterasi Dapat Dihentikan.
Herlinda Mawarni 2022	Implementasi Algoritma K-Means Untuk Segmentasi Pelanggan Pada	Algoritma <i>K-Means Clustering</i>	Data Transaksi Penjualan Sejak Tahun 2021 Sampai Desember 2022.	Cluster 1 Atau Segmen Silver, Cluster 2 Atau Segmen Gold Dan Cluster 3 Segmen Platinum Dari

	Pt. Bintang Multi Sarana Cabang Tugumulyo			Penelitian Ini Diketahui Karakteristik Setiap Segmen Pelanggan Yang Terbentuk Dari Proses Clustering Sehingga Dapat Digunakan Perusahaan Untuk Membantu Mengidentifikasi Segmen Pelanggan Mana Yang Akan Menjadi Target Pemasaran Perusahaan
Sahara Lubis 2023	Customer Relationship Management Dalam Meningkatkan Loyalitas Pelanggan Pada Doorsmeer Keluarga Nasution Menggunakan Metode Algoritma K- Means	<i>Algoritma K- means</i>	data jumlah kunjungan dan total transaksi pada usaha Doorsmeer Keluarga Nasution	a pengunjung baru ataupun yang belum menenal Doorsmeer Keluarga Nasution baik yang berada sekitar lokasi usaha maupun yang berada jauh akan mudah untuk menenal Doorsmeer Keluarga Nasution melalui pencarian pada internet sehingga meningkatkan jumlah pelanggan yang bertransaksi dengan usaha Doorsmeer Keluarga Nasution

Wicaksono 2020	Segmentasi Pelanggan Bisnis Dengan Multi Kriteria Menggunakan K-Means	<i>Algoritma K- Means</i>	Data Yang Diperoleh Dari Data Member Toko Retail Yang Melakukan Transaksi	Dalam Menerapkan Multi-Kriteria Pada Segmentasi Pelanggan Dapat Lebih Baik Dibandingkan Dengan Hanya Kriteria RFM. Sehingga Tidak Akan Salah Untuk Memperlakukan Pelanggan Sesuai Dengan Kelompok Yang Sudah Dibentuk.
Ni Putu Viona Viandari 2022	Pemetaan Pelanggan dengan LRFM dan Two Stage Clustering untuk Memenuhi Strategi Pengelolaan	<i>k-mean clustering</i>	Data yang digunakan pada penelitian dikumpulkan dari perusahaan PT. Nadia Kencana. Data yang diperoleh merupakan data transaksi periode Januari 2017 hingga Desember 2018.	Hasil perhitungan standar deviasi dari 5 cluster dengan masing – masing atribut LRFM didapatkan bahwa cluster 0, cluster 2, dan cluster 4 termasuk kelompok pelanggan new customers dengan jenis uncertain new customers (L↓R↓F↓M↓) dan strategi marketing yang diusulkan yaitu let-go strategi. Cluster 1 dan cluster 3 termasuk kelompok pelanggan potential customers dengan jenis potential loyal

				customers ($L \uparrow R \uparrow F \uparrow M \uparrow$) dan strategi marketing yang diusulkan yaitu enforced strategi.
Yosia Oktavian Pailan 2022	Segmentasi Loyalitas Pelanggan Berdasarkan RFM (Recency, Frequency, Monetary) Menggunakan K-Means pada PD. Persada Ikan	metode <i>K- Means</i>	data transaksi	hasil keakurasian data mencapai 0.770 mendekati angka 1. Saran untuk sistem ini adalah dengan mencoba menggunakan metode lain untuk mengukur dan mendapatkan hasil lebih baik.
Febriani 2020	Segmentasi Konsumen Berdasarkan Model Recency, Frequency, Monetary dengan Metode K-Means	<i>k-means clustering</i>	Segmentasi Konsumen Berdasarkan Model Recency, Frequency, Monetary dengan Metode K-Means	Klaster yang termasuk dalam segmen pelanggan potensial adalah klaster 1. Klaster tersebut perlu dipertahankan agar pelanggan tersebut tetap memiliki nilai R yang rendah, F yang tinggi dan M yang tinggi. Sementara itu, klaster yang perlu ditingkatkan lagi strateginya adalah klaster 2.
Hasnah Vithon	Pengukuran Tingkat	<i>K-means Clustering</i>	Peneliti dengan menggunakan	Perhitungan K-Means pada iterasi 1 dan

Carelsa 2023	Kepuasan Mahasiswa Terhadap Pelayanan di Kantin Kampus Menggunakan Algoritma K- means Clustering		dataset sebanyak 24 kuesioner.	iterasi 2 tidak menunjukkan adanya cluster yang mengalami perpindahan dan sampai tahap ini clustering data sudah stabil dan konvergen
Mawaddah Harahap 2022	Analisis Pemasaran Bisnis dengan Data Science : Segmentasi Kepribadian Pelanggan berdasarkan Algoritma K- Means Clustering	<i>Algoritma K- Means Clustering</i>	Kumpulan data ini dikelompokkan menjadi empat bagian yang dijadikan satu, yaitu data pelanggan, data produk, data promosi dan data tempat	80% pelanggan dikelompok ini adalah orang tua. Cluster perak adalah pelanggan berpenghasilan menengah dengan transaksi pembelian memiliki frekuensi menengah dan kebanyakan orang tua. Terakhir cluster perunggu merupakan kelompok pelanggan yang berpenghasilan rendah, sebagian sarjana dengan jumlah pembelian paling rendah, namun paling banyak mengeluh.
Satria Ardi Perdana 2024	Analisis segmentasi pelanggan menggunakan k-	<i>Algoritma K- means Clustering</i>	Penelitian ini menggunakan data riwayat transaksi	Segmentasi pelanggan menggunakan metode clustering K-Means bekerja dengan baik

	means clustering studi kasus aplikasi alfagift		pembelian pelanggan pada aplikasi belanja online Alfagift	dalam melakukan clustering pelanggan. Terbentuknya clustering pelanggan ditentukan dengan tiga kriteria yaitu umur, jenis kelamin, frekuensi pesanan, tipe pembayaran dan kota pembelian barang selama satu bulan.
Hesti Astuti 2022	Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Pengelompokan Data Pelanggan (Studi Kasus : PT. Pinus Merah Abadi)	<i>k-means clustering</i>	Data yang digunakan dalam penelitian ini berasal dari data transaksi penjualan bulan Januari-September tahun 2017. Format data yang digunakan adalah format .xlsx (Excel Document).	menghasilkan tiga cluster pelanggan dengan karakteristik pelanggan dari tiap-tiap cluster. Cluster 1 berjumlah 679 dengan kriteria recency rendah frequency rendah monetary rendah termasuk kedalam pelanggan baru, cluster 2 berjumlah 17 dengan kriteria recency tinggi frequency tinggi monetary tinggi termasuk ke dalam pelanggan yang loyal dan cluster 3 berjumlah 334 dengan kriteria recency tinggi frequency rendah

				monetary rendah termasuk kedalam pelanggan tidak loyal yang berpotensi tidak aktif .
Satria Ardi Perdana(2022)	Analisis segmentasi pelanggan menggunakan k-means clustering studi kasus aplikasi alfagift	<i>K-Means Clustering</i>	data histori transaksi pelanggan pengguna aplikasi Alfagift selama bulan Juni 2021 berdasarkan lima kategori yaitu umur, jenis kelamin, frekuensi pembelian, tipe pembayaran dan kota.	Segmentasi pelanggan menggunakan metode clustering K-Means bekerja dengan baik dalam melakukan clustering pelanggan. Terbentuknya clustering pelanggan ditentukan dengan tiga kriteria yaitu umur, jenis kelamin, frekuensi pesanan, tipe pembayaran dan kota pembelian barang selama satu bulan. Setelah dilakukannya proses segmentasi data histori transaksi pembelian pelanggan, terbentuk tiga cluster yaitu cluster 1, cluster 2 dan cluster 3. Dari hasil segmentasi pelanggan yang telah divisualisasikan dan diinterpretasikan, diharapkan dapat membantu perusahaan

				dalam memahami kebutuhan pelanggan secara lebih akurat, serta dapat digunakan sebagai data dalam menentukan strategi pemasaran yang tepat.
Apip Pramudiansyah(2021)	Segmentasi pelanggan menggunakan algoritma kmeans berdasarkan model recency frequency monetary	<i>K-Means Clustering</i>	Data yang digunakan dalam penelitian ini adalah data transaksi penjualan dengan periode satu tahun terakhir.	Berdasarkan hasil dari penelitian mengenai Segmentasi Pelanggan Menggunakan Algoritma Data Mining Berdasarkan Model Recency Frequency Monetary maka didapat kesimpulan bahwa hasil pengujian menggunakan Elbow Method didapatkan titik siku terbaik yang terbentuk diantara K3 dan K5, sehingga dapat disimpulkan bahwa jumlah klaster optimal yaitu sebanyak 4 klaster. Kemudian Algoritma yang tepat digunakan untuk model dataset yang dimiliki adalah algoritma k-means dengan nilai DBI

				sebesar 0.525. Setelah dilakukan perhitungan menggunakan algoritma kmeans dengan jumlah 4 klaster maka dihasilkan sebanyak 38,4% atau berjumlah 557 pelanggan masuk kedalam kelompok kategori Platinum, 12,6% atau berjumlah 184 pelanggan kategori Gold, 22,7% atau berjumlah 330 pelanggan kategori Silver, dan 26.% atau berjumlah 378 pelanggan kategori Bronze
Ahmad Pudoli(2024)	Implementasi Algoritma K-Means Clustering dalam Pengelompokan Produk Pigeon Pada PT. Digital Niaga Indonesia Berdasarkan Analisis Recency, Frequency,	<i>K-Means Clustering</i>	Data yang akan dipakai untuk penelitian atau riset merupakan nilai data pada bulan Agustus hingga November 2022. Dan tanggal penelitian 7 Januari 2023.	a. Evauasi Cluster menggunakan perhitungan DBI menghasilkan 0.520 yang dimana nilai ini bagus karena jika nilai mendekati 0 maka hasil yang diperoleh akan semakin bagus. b. Pengujian aplikasi didapatkan 557 produk dengan hasil akhir

	Monetary (RFM)			akhir 3 Cluster yaitu Cluster 1 (C1) 100 data, Cluster 2 (C2) sebanyak 20 data, dan Cluster 3 (C3) sebanyak 437 data dari 124473 data transaksi. c. Metode RFM dapat menentukan bahwa Cluster 1 adalah produk yang kurang laris, Cluster 2 paling laris, dan Cluster 3 laris. d. Dengan adanya pengelompokan data ini diharapkan, manajemen dan tim marketing PT. Digital Niaga Indonesia dapat mengetahui barang yang laris, laris, kurang laris oleh pembeli di setiap marketplace dengan mudah.
--	-------------------	--	--	---