

Penerapan Data Mining untuk Customer Intelligence dalam Meningkatkan Strategi Bisnis Retail Menggunakan K-Means dan FP-Growth

Vivi Ananda Putri¹, Faiz Perdana², Arael Idfi Krismarwazie³

¹²³Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pamulang, Tangerang Selatan, Indonesia

¹Vivianaputrivn25@gmail.com, ²faiz.p2018@gmail.com, ³aurael.idfi1611@gmail.com

Abstract

The increasingly competitive retail industry requires companies to gain a deeper understanding of consumer characteristics and behavior. Leveraging transaction data stored in corporate databases can serve as a source of strategic information to support business decision-making. This study aims to apply data mining techniques to build customer intelligence, thereby enhancing the effectiveness of retail business strategies. The methodology employs the K-Means algorithm for customer segmentation based on transaction characteristics and the FP-Growth algorithm to identify patterns of products frequently purchased together. The research process begins with the collection and cleaning of transaction data, followed by data transformation, customer clustering using K-Means, and product association analysis using FP-Growth. The customer segmentation results are expected to identify consumer groups based on activity levels and transaction values, enabling the company to formulate more targeted marketing strategies. Meanwhile, the product association patterns derived from FP-Growth can be utilized to design promotions, determine product placement, and develop cross-selling strategies. By combining these two algorithms, this study aims to generate valuable insights that help the company understand customer behavior, boost consumer loyalty, and sustainably optimize retail business performance.

Keywords: Data Mining, Customer Intelligence, Retail, K-Means, FP-Growth, Customer Segmentation, Association Analysis.

Abstrak

Perkembangan industri ritel yang semakin kompetitif menuntut perusahaan untuk memahami karakteristik serta perilaku konsumennya secara lebih mendalam. Pemanfaatan data transaksi yang tersimpan dalam basis data perusahaan dapat menjadi sumber informasi strategis dalam mendukung pengambilan keputusan bisnis. Penelitian ini bertujuan menerapkan teknik data mining untuk membangun customer intelligence guna meningkatkan efektivitas strategi bisnis ritel. Metode yang digunakan adalah algoritma K-Means untuk melakukan segmentasi pelanggan berdasarkan karakteristik transaksi dan algoritma FP-Growth untuk menemukan pola keterkaitan pembelian produk yang sering terjadi secara bersamaan. Proses penelitian diawali dengan pengumpulan dan pembersihan data transaksi, dilanjutkan dengan transformasi data, pengelompokan pelanggan menggunakan K-Means, serta analisis asosiasi produk menggunakan FP-Growth. Hasil segmentasi pelanggan diharapkan mampu mengidentifikasi kelompok konsumen berdasarkan tingkat aktivitas dan nilai transaksi sehingga perusahaan dapat menyusun strategi pemasaran yang lebih tepat sasaran. Sementara itu, pola asosiasi produk yang dihasilkan dari FP-Growth dapat dimanfaatkan dalam penyusunan promosi, penempatan produk, dan strategi cross-selling. Melalui kombinasi kedua algoritma tersebut, penelitian ini diharapkan mampu menghasilkan informasi yang bernilai bagi perusahaan dalam memahami perilaku pelanggan, meningkatkan loyalitas konsumen, serta mengoptimalkan kinerja bisnis ritel secara berkelanjutan.

Kata Kunci: Data Mining, Customer Intelligence, Retail, K-Means, FP-Growth, Segmentasi Pelanggan, Analisis Asosiasi.

A. PENDAHULUAN

Transformasi digital yang berlangsung secara masif telah mengubah pola persaingan pada sektor ritel. Perusahaan tidak lagi hanya berkompetisi melalui kualitas produk dan harga, tetapi juga melalui kemampuan dalam memahami karakteristik serta perilaku konsumennya. Aktivitas transaksi yang terjadi setiap hari menghasilkan akumulasi

data dalam jumlah besar yang menyimpan berbagai informasi bernilai bagi organisasi bisnis. Apabila data tersebut hanya disimpan sebagai arsip operasional, maka potensi pengetahuan yang terkandung di dalamnya tidak dapat dimanfaatkan secara optimal. Oleh sebab itu, diperlukan pendekatan analitis yang mampu mengekstraksi informasi tersembunyi dari data transaksi sehingga dapat digunakan sebagai dasar dalam penyusunan strategi

pemasaran yang lebih efektif dan berorientasi pada kebutuhan pelanggan.

Perkembangan konsep customer intelligence menjadi salah satu pendekatan yang banyak diterapkan oleh perusahaan modern dalam memahami perilaku konsumen secara lebih mendalam. Customer intelligence memungkinkan organisasi memperoleh wawasan mengenai kebiasaan pembelian, preferensi produk, tingkat loyalitas, serta potensi nilai pelanggan bagi perusahaan. Informasi tersebut dapat dimanfaatkan untuk menyusun program promosi yang lebih terarah, meningkatkan kualitas layanan, serta memperkuat hubungan jangka panjang dengan konsumen. Dalam lingkungan bisnis ritel yang dinamis, pemanfaatan customer intelligence menjadi faktor penting karena mampu membantu perusahaan mengidentifikasi peluang pasar sekaligus mempertahankan pelanggan yang memberikan kontribusi terbesar terhadap pendapatan perusahaan.

Salah satu cabang ilmu yang banyak digunakan dalam proses penggalian informasi dari kumpulan data berukuran besar adalah data mining. Teknik ini memungkinkan proses identifikasi pola, kecenderungan, hubungan, maupun karakteristik tertentu yang sebelumnya sulit diketahui melalui pengamatan manual. Berbagai penelitian menunjukkan bahwa data mining mampu menghasilkan informasi strategis yang dapat mendukung proses pengambilan keputusan bisnis secara lebih akurat dan berbasis data. Dalam konteks pengelolaan pelanggan, penerapan data mining dapat membantu perusahaan mengenali kelompok konsumen yang memiliki karakteristik serupa sehingga strategi pemasaran dapat disesuaikan dengan kebutuhan masing-masing segmen pelanggan.

Algoritma K-Means merupakan salah satu metode clustering yang banyak digunakan dalam segmentasi pelanggan karena memiliki mekanisme pengelompokan yang relatif sederhana, efisien, dan mampu mengolah data dalam jumlah besar. Melalui proses pengelompokan tersebut, pelanggan dapat diklasifikasikan berdasarkan kesamaan karakteristik transaksi sehingga perusahaan dapat mengetahui kelompok pelanggan yang memiliki nilai tinggi, pelanggan potensial, maupun pelanggan yang memerlukan perhatian khusus. Sejumlah penelitian terdahulu menunjukkan bahwa penerapan K-Means mampu menghasilkan segmentasi pelanggan yang mendukung peningkatan efektivitas strategi Customer Relationship Management (CRM), pengembangan program loyalitas, serta perencanaan promosi yang lebih tepat sasaran.

Selain memahami karakteristik pelanggan, perusahaan ritel juga perlu mengetahui keterkaitan antarproduk yang sering dibeli secara bersamaan oleh konsumen. Informasi tersebut dapat diperoleh melalui teknik association rule mining menggunakan algoritma FP-Growth. Metode ini mampu menemukan pola kombinasi produk yang memiliki tingkat kemunculan tinggi dalam data transaksi tanpa

menghasilkan proses pencarian yang terlalu kompleks. Hasil analisis FP-Growth dapat dimanfaatkan untuk mendukung strategi cross-selling, penataan tata letak produk, penyusunan paket promosi, serta rekomendasi produk yang sesuai dengan perilaku pembelian pelanggan. Dengan demikian, perusahaan dapat meningkatkan peluang penjualan sekaligus memberikan pengalaman berbelanja yang lebih relevan bagi konsumen.

Berdasarkan uraian tersebut, penelitian ini bertujuan menerapkan teknik data mining untuk membangun customer intelligence dalam mendukung peningkatan strategi bisnis ritel melalui kombinasi algoritma K-Means dan FP-Growth. Algoritma K-Means digunakan untuk mengelompokkan pelanggan berdasarkan karakteristik transaksi yang dimiliki, sedangkan FP-Growth dimanfaatkan untuk mengidentifikasi pola keterkaitan produk yang sering muncul dalam aktivitas pembelian. Integrasi kedua metode tersebut diharapkan mampu menghasilkan informasi yang lebih komprehensif mengenai perilaku pelanggan dan pola konsumsi sehingga dapat membantu perusahaan menyusun strategi pemasaran yang lebih efektif, meningkatkan loyalitas konsumen, serta mengoptimalkan kinerja bisnis secara berkelanjutan.

B. METODE

Dalam proses segmentasi pelanggan menggunakan algoritma K-Means, salah satu tahapan yang memiliki peranan penting adalah penentuan jumlah cluster yang optimal. Pemilihan jumlah kelompok yang kurang tepat berpotensi menghasilkan segmentasi yang tidak representatif sehingga informasi yang diperoleh menjadi kurang akurat. Oleh karena itu, diperlukan suatu pendekatan yang mampu membantu peneliti menentukan jumlah cluster terbaik berdasarkan karakteristik data yang dianalisis. Salah satu teknik yang paling banyak digunakan untuk tujuan tersebut adalah Elbow Method.

Elbow Method merupakan pendekatan evaluatif yang digunakan untuk mengidentifikasi jumlah cluster yang paling sesuai pada algoritma K-Means. Teknik ini bekerja dengan menghitung nilai Within Cluster Sum of Squares (WCSS) atau jumlah kuadrat jarak antaranggota data terhadap pusat cluster masing-masing. Semakin kecil nilai WCSS yang diperoleh, maka semakin tinggi tingkat homogenitas data dalam suatu kelompok. Namun demikian, penambahan jumlah cluster secara terus-menerus tidak selalu menghasilkan peningkatan kualitas segmentasi yang signifikan.

Prinsip dasar Elbow Method adalah mencari titik perubahan yang menyerupai bentuk siku atau elbow pada grafik hubungan antara jumlah cluster dan nilai WCSS. Pada tahap awal, penambahan jumlah cluster biasanya menyebabkan penurunan nilai WCSS yang cukup besar. Akan tetapi, setelah mencapai titik tertentu, penurunan tersebut cenderung melambat. Titik perlambatan inilah yang dianggap sebagai jumlah cluster yang paling optimal

karena mampu menghasilkan keseimbangan antara kompleksitas model dan kualitas pengelompokan data.

Dalam penelitian customer intelligence pada sektor retail, Elbow Method memiliki fungsi strategis karena membantu menghasilkan segmentasi pelanggan yang lebih akurat. Melalui penentuan jumlah cluster yang tepat, perusahaan dapat mengidentifikasi kelompok pelanggan berdasarkan pola transaksi, frekuensi pembelian, nilai pengeluaran, maupun tingkat loyalitas secara lebih efektif. Segmentasi yang baik akan menghasilkan informasi yang lebih relevan untuk mendukung penyusunan strategi pemasaran yang berbasis data.

Penerapan Elbow Method pada penelitian ini diawali dengan proses pembersihan data transaksi pelanggan. Setelah data dinyatakan layak untuk dianalisis, algoritma K-Means dijalankan secara berulang menggunakan berbagai jumlah cluster, misalnya mulai dari dua hingga sepuluh kelompok. Pada setiap percobaan, sistem akan menghitung nilai WCSS yang kemudian digunakan sebagai dasar dalam pembentukan grafik Elbow. Hasil visualisasi tersebut menjadi acuan dalam menentukan jumlah cluster terbaik yang akan digunakan pada tahap segmentasi pelanggan.

Keunggulan utama Elbow Method terletak pada kemampuannya memberikan gambaran visual yang mudah dipahami oleh peneliti maupun pengambil keputusan. Pendekatan ini tidak memerlukan perhitungan yang terlalu kompleks sehingga dapat diterapkan pada berbagai jenis data transaksi retail. Selain itu, metode ini juga banyak digunakan dalam penelitian data mining karena mampu memberikan hasil yang cukup konsisten dalam menentukan jumlah cluster yang efisien dan representatif.

Dalam konteks customer intelligence, hasil segmentasi yang diperoleh dari kombinasi Elbow Method dan K-Means dapat digunakan untuk mengenali karakteristik pelanggan secara lebih rinci. Kelompok pelanggan dengan tingkat transaksi tinggi dapat diprioritaskan sebagai target program loyalitas, sedangkan kelompok pelanggan dengan aktivitas pembelian rendah dapat menjadi sasaran strategi promosi khusus. Dengan demikian, perusahaan dapat mengalokasikan sumber daya pemasaran secara lebih efektif dan terarah.

Selain mendukung segmentasi pelanggan, hasil pengelompokan yang diperoleh juga menjadi fondasi bagi proses analisis berikutnya menggunakan algoritma FP-Growth. Setiap cluster pelanggan dapat dianalisis untuk mengetahui pola kombinasi produk yang sering dibeli secara bersamaan. Pendekatan ini memungkinkan perusahaan memperoleh pemahaman yang lebih komprehensif mengenai perilaku konsumsi pada masing-masing kelompok pelanggan sehingga strategi cross-selling dan up-selling dapat dirancang secara lebih optimal.

Beberapa penelitian di Indonesia menunjukkan bahwa penggunaan Elbow Method mampu meningkatkan kualitas

hasil clustering karena jumlah cluster yang digunakan tidak ditentukan secara subjektif. Pendekatan tersebut membantu peneliti memperoleh kelompok data yang lebih stabil dan memiliki tingkat kemiripan internal yang tinggi. Oleh sebab itu, Elbow Method sering dijadikan sebagai prosedur pendukung sebelum penerapan algoritma K-Means pada berbagai bidang, termasuk analisis pelanggan, pemasaran digital, manajemen hubungan pelanggan, dan pengembangan strategi bisnis.

Berdasarkan uraian tersebut, dapat disimpulkan bahwa Elbow Method merupakan teknik yang efektif untuk menentukan jumlah cluster optimal pada algoritma K-Means. Dalam penelitian mengenai penerapan data mining untuk customer intelligence pada bisnis retail, metode ini berperan penting dalam menghasilkan segmentasi pelanggan yang lebih akurat sehingga mampu mendukung pengambilan keputusan strategis. Kombinasi Elbow Method, K-Means, dan FP-Growth diharapkan dapat menghasilkan informasi yang bernilai dalam memahami perilaku pelanggan, meningkatkan kualitas layanan, serta memperkuat daya saing perusahaan retail secara berkelanjutan.

C. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan serangkaian informasi analitis yang dimanfaatkan untuk mendukung pengembangan customer intelligence pada sektor retail. Proses pengolahan data dilakukan terhadap data transaksi pelanggan yang mencakup frekuensi pembelian, jumlah produk yang diperoleh, nilai transaksi, serta pola kombinasi produk yang muncul pada setiap aktivitas pembelian. Tahapan analisis memanfaatkan algoritma K-Means untuk melakukan segmentasi pelanggan dan algoritma FP-Growth untuk mengidentifikasi hubungan antarproduk yang sering dibeli secara bersamaan. Integrasi kedua pendekatan tersebut menghasilkan wawasan yang dapat digunakan sebagai dasar dalam penyusunan strategi bisnis yang lebih terukur dan berorientasi pada kebutuhan pelanggan.

Tahap awal penelitian difokuskan pada penentuan jumlah cluster yang optimal menggunakan Elbow Method. Berdasarkan hasil pengujian beberapa nilai cluster, diperoleh titik penurunan nilai Within Cluster Sum of Squares (WCSS) yang mulai melandai pada jumlah cluster tertentu sehingga dianggap sebagai jumlah kelompok yang paling representatif. Penggunaan Elbow Method membantu mengurangi subjektivitas dalam menentukan jumlah cluster sehingga hasil segmentasi menjadi lebih stabil dan mampu merepresentasikan karakteristik pelanggan secara lebih akurat. Temuan ini sejalan dengan penelitian Maori dan Evanita (2023) yang menjelaskan bahwa Elbow Method dapat membantu memperoleh jumlah cluster yang optimal pada proses K-Means Clustering.

Setelah jumlah cluster ditentukan, proses pengelompokan pelanggan menggunakan algoritma K-Means menghasilkan beberapa segmen pelanggan dengan karakteristik yang berbeda. Sebagian besar pelanggan

berada pada kelompok dengan intensitas transaksi menengah, sedangkan sebagian lainnya termasuk dalam kelompok pelanggan bernilai tinggi yang memiliki frekuensi pembelian lebih besar dan nilai transaksi yang relatif dominan. Keberadaan kelompok pelanggan bernilai tinggi menunjukkan adanya segmen yang memberikan kontribusi signifikan terhadap pendapatan perusahaan sehingga perlu mendapatkan perhatian khusus dalam strategi pemasaran dan program loyalitas. Hasil ini sejalan dengan penelitian Fazri, Muhammad, dan Purnama (2023) yang menunjukkan bahwa K-Means mampu mengelompokkan pelanggan berdasarkan karakteristik transaksi sehingga memudahkan perusahaan dalam mengidentifikasi pelanggan potensial.

Hasil segmentasi juga memperlihatkan adanya kelompok pelanggan dengan tingkat aktivitas transaksi yang rendah. Meskipun kontribusi kelompok ini tidak sebesar pelanggan utama, keberadaannya tetap memiliki nilai strategis karena berpotensi ditingkatkan melalui pendekatan pemasaran yang lebih personal. Perusahaan dapat memanfaatkan informasi tersebut untuk menyusun promosi khusus, pemberian diskon berkala, maupun program retensi pelanggan yang bertujuan meningkatkan frekuensi pembelian. Dengan demikian, strategi pemasaran tidak hanya berfokus pada pelanggan terbaik, tetapi juga mampu mengembangkan pelanggan yang masih memiliki potensi pertumbuhan.

Analisis customer intelligence yang diperoleh dari hasil clustering menunjukkan bahwa setiap kelompok pelanggan memiliki pola perilaku yang berbeda. Pelanggan dengan aktivitas transaksi tinggi cenderung memiliki loyalitas yang lebih baik dibandingkan pelanggan dengan frekuensi pembelian rendah. Informasi tersebut menjadi dasar bagi perusahaan dalam merancang strategi Customer Relationship Management (CRM) yang lebih efektif. Temuan ini mendukung penelitian Tivano, Ruslianto, dan Prawira (2023) yang menyatakan bahwa segmentasi pelanggan dapat membantu perusahaan memahami karakteristik konsumen secara lebih rinci sehingga strategi pemasaran dapat diarahkan sesuai kebutuhan masing-masing segmen.

Selain melakukan segmentasi pelanggan, penelitian ini juga menerapkan algoritma FP-Growth untuk menemukan pola asosiasi produk dari data transaksi retail. Hasil pengolahan menunjukkan adanya beberapa kombinasi produk yang memiliki tingkat kemunculan bersamaan cukup tinggi. Pola tersebut mengindikasikan bahwa pelanggan yang membeli suatu produk tertentu memiliki kecenderungan untuk membeli produk lain dalam transaksi yang sama. Informasi ini menjadi sumber pengetahuan yang penting dalam mendukung pengambilan keputusan bisnis berbasis data.

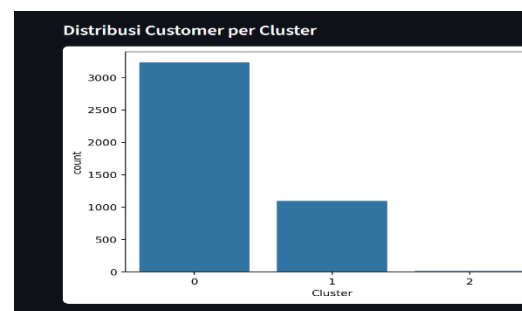
Pola asosiasi yang dihasilkan oleh FP-Growth dapat dimanfaatkan untuk meningkatkan efektivitas strategi penjualan. Produk yang sering muncul secara bersamaan dapat ditempatkan pada area yang berdekatan, dijadikan

paket promosi, atau digunakan sebagai dasar rekomendasi produk kepada pelanggan. Pendekatan tersebut berpotensi meningkatkan nilai transaksi rata-rata sekaligus mendorong terjadinya pembelian tambahan. Dengan memanfaatkan informasi asosiasi produk, perusahaan dapat mengoptimalkan peluang cross-selling dan up-selling secara lebih terarah.

Kombinasi hasil segmentasi pelanggan dan analisis asosiasi produk memberikan gambaran yang lebih komprehensif mengenai perilaku konsumen. Perusahaan tidak hanya mengetahui siapa pelanggan yang paling bernilai, tetapi juga memahami produk apa saja yang cenderung dibeli oleh masing-masing kelompok pelanggan. Informasi ini memungkinkan perusahaan merancang strategi pemasaran yang lebih presisi, mulai dari penentuan target promosi hingga penyusunan rekomendasi produk yang relevan berdasarkan karakteristik pelanggan. Pendekatan customer intelligence semacam ini mampu meningkatkan efektivitas pengambilan keputusan bisnis karena seluruh strategi didasarkan pada hasil analisis data yang objektif.

Secara keseluruhan, hasil penelitian menunjukkan bahwa penerapan data mining menggunakan kombinasi Elbow Method, K-Means, dan FP-Growth mampu menghasilkan informasi strategis yang mendukung peningkatan kinerja bisnis retail. Segmentasi pelanggan memberikan pemahaman mengenai karakteristik konsumen berdasarkan perilaku transaksi, sedangkan analisis asosiasi produk membantu mengidentifikasi pola pembelian yang dapat dimanfaatkan untuk meningkatkan penjualan. Integrasi kedua metode tersebut menghasilkan customer intelligence yang lebih komprehensif sehingga perusahaan dapat menyusun strategi pemasaran yang lebih efektif, meningkatkan loyalitas pelanggan, serta memperkuat daya saing bisnis dalam jangka panjang.

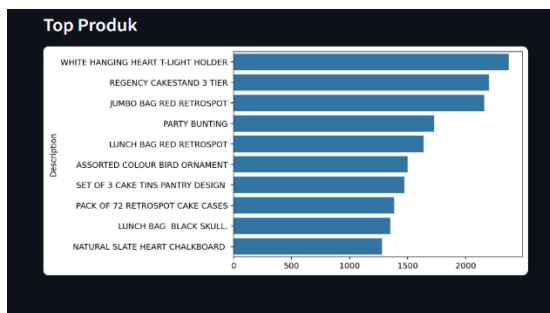
Gambar



Gambar 1. Distribusi Customer

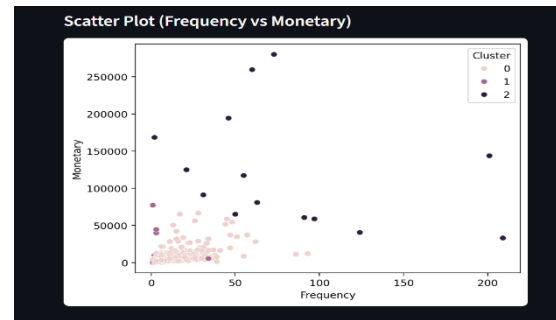
Berdasarkan grafik Distribusi Customer per Cluster, terlihat bahwa hasil pengelompokan pelanggan menggunakan algoritma K-Means menghasilkan tiga kelompok dengan jumlah anggota yang berbeda secara signifikan. Cluster 0 mendominasi distribusi data dengan jumlah pelanggan terbanyak, yaitu lebih dari 3.000 pelanggan, yang menunjukkan bahwa mayoritas konsumen memiliki karakteristik transaksi yang relatif serupa dan berada pada segmen utama perusahaan. Sementara itu,

Cluster 1 memiliki jumlah pelanggan yang jauh lebih sedikit, sekitar 1.000 pelanggan, sehingga dapat diinterpretasikan sebagai kelompok konsumen dengan pola aktivitas pembelian yang berbeda dari segmen dominan. Di sisi lain, Cluster 2 hanya berisi sejumlah kecil pelanggan, yang mengindikasikan adanya segmen khusus dengan karakteristik transaksi yang unik atau tidak umum dibandingkan kelompok lainnya. Ketimpangan sebaran pelanggan pada setiap cluster ini menunjukkan bahwa perilaku konsumen dalam bisnis retail tidak terdistribusi secara merata, sehingga perusahaan perlu menerapkan strategi pemasaran, program loyalitas, dan pendekatan layanan yang disesuaikan dengan karakteristik masing-masing segmen pelanggan guna meningkatkan efektivitas customer intelligence dan mendukung pengambilan keputusan bisnis yang lebih tepat sasaran.



Gambar 2. Top Produk

Berdasarkan grafik Top Produk, terlihat bahwa beberapa produk memiliki tingkat permintaan yang jauh lebih tinggi dibandingkan produk lainnya. Produk WHITE HANGING HEART T-LIGHT HOLDER menempati posisi teratas sebagai barang yang paling sering dibeli oleh pelanggan, diikuti oleh REGENCY CAKESTAND 3 TIER dan JUMBO BAG RED RETROSPOT yang juga menunjukkan volume penjualan yang relatif besar. Dominasi beberapa produk tersebut mengindikasikan adanya preferensi konsumen yang kuat terhadap jenis barang tertentu, sehingga produk-produk tersebut dapat dikategorikan sebagai produk unggulan yang memberikan kontribusi signifikan terhadap total transaksi perusahaan. Sementara itu, produk lain pada daftar sepuluh besar tetap menunjukkan tingkat penjualan yang cukup tinggi meskipun berada di bawah produk utama. Temuan ini memberikan informasi berharga bagi perusahaan retail dalam menentukan prioritas pengelolaan persediaan, menyusun strategi promosi, serta merancang rekomendasi produk yang lebih relevan berdasarkan minat pelanggan. Dengan memahami produk yang paling diminati, perusahaan dapat meningkatkan efektivitas pengambilan keputusan bisnis sekaligus mengoptimalkan peluang peningkatan pendapatan melalui pemanfaatan pola pembelian konsumen yang telah teridentifikasi.



Gambar 3. Scatter Plot

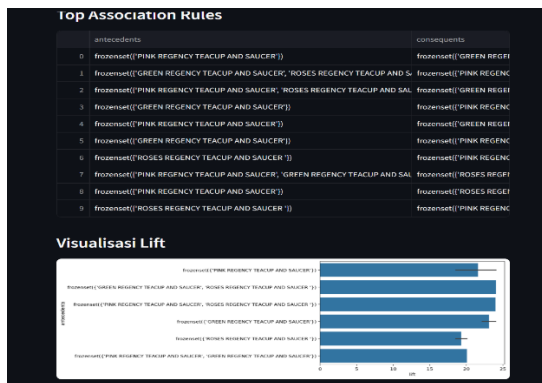
Berdasarkan grafik Scatter Plot (Frequency vs Monetary), terlihat bahwa hasil segmentasi pelanggan membentuk beberapa kelompok dengan karakteristik transaksi yang berbeda. Sebagian besar pelanggan terkonsentrasi pada area dengan nilai frekuensi pembelian dan nilai transaksi yang relatif rendah hingga menengah, yang menunjukkan bahwa kelompok tersebut melakukan aktivitas belanja secara tidak terlalu sering dengan kontribusi pengeluaran yang masih terbatas. Di sisi lain, terdapat sejumlah pelanggan yang berada jauh dari kumpulan data utama dengan nilai frekuensi maupun nilai transaksi yang jauh lebih tinggi. Kelompok ini dapat diinterpretasikan sebagai pelanggan bernilai tinggi yang memiliki intensitas pembelian lebih besar dan memberikan kontribusi pendapatan yang signifikan bagi perusahaan. Sebaran titik yang tidak merata pada grafik juga mengindikasikan adanya perbedaan perilaku konsumen yang cukup jelas antarcluster, sehingga segmentasi yang dihasilkan mampu membantu perusahaan dalam mengenali pelanggan prioritas, menyusun strategi pemasaran yang lebih terarah, serta mengembangkan program loyalitas yang disesuaikan dengan karakteristik masing-masing segmen pelanggan untuk mendukung peningkatan kinerja bisnis retail secara berkelanjutan.

Cluster	Customer ID	Recency	Frequency	Monetary
0	15282.1578	40.9839	4.8527	2005.836
1	15354.226	246.0238	1.5828	629.3574
2	15310.6429	7.1429	80.2143	122748.7921

Gambar 4. Profil RFM per Cluster

Berdasarkan tabel Profil RFM per Cluster, terlihat adanya perbedaan karakteristik pelanggan yang cukup mencolok pada setiap kelompok hasil segmentasi. Cluster 2 menunjukkan nilai Recency paling rendah serta nilai Frequency dan Monetary paling tinggi dibandingkan kelompok lainnya, yang mengindikasikan bahwa pelanggan dalam segmen ini melakukan transaksi dalam waktu yang lebih dekat, berbelanja lebih sering, dan memberikan kontribusi nilai pembelian yang sangat besar. Kondisi tersebut menunjukkan bahwa Cluster 2 dapat dikategorikan sebagai kelompok pelanggan utama yang memiliki tingkat loyalitas dan nilai bisnis tertinggi bagi perusahaan. Sementara itu, Cluster 0 menempati posisi menengah dengan aktivitas transaksi dan nilai pembelian yang cukup baik, sehingga dapat dianggap sebagai

pelanggan potensial yang masih berpeluang untuk ditingkatkan melalui program pemasaran yang tepat. Di sisi lain, Cluster 1 memiliki nilai Recency paling tinggi serta nilai Frequency dan Monetary paling rendah, yang menunjukkan bahwa pelanggan pada kelompok ini relatif jarang melakukan transaksi dan memberikan kontribusi pendapatan yang lebih kecil. Hasil analisis RFM tersebut memberikan gambaran yang lebih mendalam mengenai perilaku konsumen sehingga perusahaan retail dapat merancang strategi customer intelligence yang lebih terarah, mulai dari mempertahankan pelanggan bernilai tinggi, mengembangkan pelanggan potensial, hingga mengaktifkan kembali pelanggan yang mulai tidak aktif agar mampu meningkatkan efektivitas strategi bisnis secara berkelanjutan.



Gambar 5. Top Association Rules

Berdasarkan hasil Top Association Rules dan Visualisasi Lift, terlihat adanya hubungan keterkaitan yang sangat kuat antara beberapa produk yang sering muncul secara bersamaan dalam satu transaksi pelanggan. Kombinasi produk seperti PINK REGENCY TEACUP AND SAUCER, GREEN REGENCY TEACUP AND SAUCER, dan ROSES REGENCY TEACUP AND SAUCER menunjukkan tingkat asosiasi yang tinggi, yang mengindikasikan bahwa pelanggan yang membeli salah satu produk tersebut memiliki kecenderungan besar untuk membeli produk lainnya dalam waktu yang sama. Nilai lift yang relatif tinggi pada setiap aturan asosiasi mengisyaratkan bahwa keterkaitan antarproduk tersebut tidak terjadi secara kebetulan, melainkan mencerminkan pola pembelian yang konsisten dari konsumen. Temuan ini memberikan wawasan penting bagi perusahaan retail dalam menyusun strategi cross-selling, pembuatan paket produk, penempatan barang pada area yang berdekatan, serta pengembangan sistem rekomendasi produk yang lebih relevan. Dengan memanfaatkan pola asosiasi yang dihasilkan oleh algoritma FP-Growth, perusahaan dapat memahami preferensi pelanggan secara lebih mendalam sehingga mampu meningkatkan peluang penjualan, memperbaiki pengalaman berbelanja, dan mendukung penerapan customer intelligence yang lebih efektif dalam pengambilan keputusan bisnis.

Pelanggan Berisiko Churn

Jumlah: 14

Customer ID	Recency	Frequency	Monetary	Cluster
55	12415	24	21	124914.53
326	12748	1	209	33053.19
562	13089	3	97	58762.08
996	13694	4	50	65039.62
1333	14156	10	55	117210.08
1689	14646	2	73	280206.02
1879	14911	1	201	143711.17
2176	15311	1	91	60632.75
2702	16029	39	63	80850.84
3008	16446	1	2	168472.5

Gambar 6. Resiko Churn Pelanggan

Berdasarkan tabel Pelanggan Berisiko Churn, teridentifikasi sebanyak 14 pelanggan yang memiliki potensi untuk mengurangi intensitas transaksi atau bahkan berhenti melakukan pembelian di masa mendatang. Meskipun sebagian pelanggan dalam kelompok ini masih menunjukkan nilai Monetary dan Frequency yang relatif tinggi, pola aktivitas mereka mengindikasikan adanya perubahan perilaku yang perlu mendapatkan perhatian khusus dari perusahaan. Kondisi tersebut menunjukkan bahwa pelanggan bernilai besar tidak selalu memiliki tingkat loyalitas yang stabil, sehingga tetap berpeluang mengalami penurunan keterlibatan apabila tidak diberikan strategi pemeliharaan hubungan yang tepat. Informasi ini menjadi bagian penting dalam penerapan customer intelligence karena memungkinkan perusahaan mengenali pelanggan yang berpotensi hilang sebelum benar-benar meninggalkan layanan. Dengan memanfaatkan hasil analisis tersebut, perusahaan retail dapat menyusun langkah antisipatif seperti pemberian penawaran personal, program penghargaan pelanggan, promosi eksklusif, maupun komunikasi yang lebih intensif untuk meningkatkan kepuasan dan mempertahankan hubungan jangka panjang. Oleh karena itu, identifikasi pelanggan berisiko churn tidak hanya berfungsi sebagai alat pemantauan perilaku konsumen, tetapi juga sebagai dasar dalam merancang strategi retensi yang lebih efektif guna menjaga stabilitas pendapatan dan meningkatkan keberlanjutan bisnis.

D. PENUTUP / KESIMPULAN

Berdasarkan keseluruhan tahapan penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan data mining mampu menghasilkan informasi yang bernilai dalam mendukung pengembangan customer intelligence pada bisnis retail. Melalui pemanfaatan data transaksi pelanggan, penelitian ini berhasil mengungkap karakteristik perilaku konsumen yang sebelumnya tersembunyi di dalam kumpulan data operasional. Proses segmentasi menggunakan algoritma K-Means yang didukung oleh Elbow Method sebagai teknik penentuan jumlah cluster optimal mampu menghasilkan pengelompokan pelanggan berdasarkan kesamaan pola transaksi. Hasil pengelompokan tersebut menunjukkan adanya perbedaan karakteristik pada setiap segmen pelanggan, mulai dari kelompok pelanggan bernilai tinggi, pelanggan potensial, hingga pelanggan dengan tingkat aktivitas transaksi yang relatif rendah. Informasi tersebut

memberikan gambaran yang lebih mendalam mengenai kondisi pelanggan sehingga perusahaan dapat menyusun strategi pemasaran yang lebih tepat sasaran.

Selain segmentasi pelanggan, penerapan algoritma FP-Growth juga berhasil mengidentifikasi pola keterkaitan antarproduk yang sering muncul secara bersamaan dalam aktivitas pembelian. Hasil analisis asosiasi menunjukkan bahwa terdapat beberapa kombinasi produk yang memiliki hubungan kuat dan berpotensi dimanfaatkan dalam strategi penjualan. Informasi tersebut dapat digunakan sebagai dasar dalam penyusunan program promosi, penawaran paket produk, rekomendasi barang, serta pengaturan tata letak produk yang lebih efektif. Temuan ini membuktikan bahwa analisis pola pembelian tidak hanya membantu perusahaan memahami preferensi konsumen, tetapi juga memberikan peluang untuk meningkatkan nilai transaksi melalui penerapan strategi cross-selling dan up-selling yang lebih terarah.

Saran

Berdasarkan temuan penelitian yang telah diperoleh, perusahaan retail disarankan untuk memanfaatkan hasil segmentasi pelanggan dan analisis pola pembelian sebagai landasan dalam penyusunan strategi bisnis yang lebih berbasis data. Informasi customer intelligence yang dihasilkan melalui penerapan algoritma K-Means dan FP-Growth dapat digunakan untuk mengembangkan program loyalitas pelanggan, menyusun promosi yang lebih terarah, serta meningkatkan efektivitas pengelolaan produk sesuai dengan karakteristik masing-masing segmen konsumen. Selain itu, perusahaan juga dianjurkan untuk melakukan pembaruan dan pengelolaan data transaksi secara berkelanjutan agar informasi yang diperoleh tetap relevan dengan perubahan perilaku pelanggan dan dinamika pasar yang terus berkembang. Dengan pemanfaatan data yang lebih optimal, perusahaan diharapkan mampu meningkatkan kualitas pelayanan, memperkuat hubungan dengan pelanggan, serta mendorong pertumbuhan penjualan secara berkesinambungan.

Untuk penelitian selanjutnya, disarankan agar cakupan analisis diperluas dengan menambahkan variabel maupun sumber data lain yang dapat memberikan gambaran perilaku pelanggan secara lebih menyeluruh. Pengembangan penelitian juga dapat dilakukan dengan mengombinasikan metode K-Means dan FP-Growth dengan algoritma lain, seperti Decision Tree, Random Forest, ataupun teknik pembelajaran mesin yang lebih modern guna memperoleh hasil analisis yang lebih mendalam. Selain itu, penelitian berikutnya dapat mengintegrasikan data dari berbagai kanal digital, seperti aktivitas pelanggan pada aplikasi, media sosial, maupun platform perdagangan elektronik sehingga proses

pembentukan customer intelligence menjadi lebih komprehensif. Melalui pengembangan tersebut, hasil penelitian diharapkan mampu menghasilkan rekomendasi strategi bisnis yang lebih akurat, adaptif, dan sesuai dengan kebutuhan perusahaan retail dalam menghadapi persaingan bisnis yang semakin kompetitif.

E. DAFTAR PUSTAKA

- Fazri, A. H., Muhammad, A., & Purnama, P. A. W. (2023). Algoritma K-Means Cluster untuk Segmentasi Pelanggan. *Computer Based Information System Journal*, 11(2), 42–51.
- Maori, N. A., & Evanita, E. (2023). Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering. *Simetris: Jurnal Teknik Mesin, Industri, Elektro dan Ilmu Komputer*, 14(2), 1419–1428.
- Tivano, D., Ruslianto, I., & Prawira, D. (2023). Sistem Segmentasi Pelanggan Potensial Menggunakan Metode K-Means dan Analisis RFM. *Coding: Jurnal Komputer dan Aplikasi*, 11(3).
- Muhammad, H., Haryadi, H., & Sularno, S. (2024). Segmentasi Pelanggan Berdasarkan Pendapatan dan Skor Pengeluaran Menggunakan Metode K-Means. *Jurnal Sains dan Teknologi (JSIT)*, 4(3), 325–337.
- Fikri, A., Hutabarat, B. F., & Khaira, U. (2023). Komparasi K-Means Clustering dan Complete Linkage dalam Pengelompokan Penyaluran Pinjaman Financial Technology. *Jurnal Ilmiah Media Sisfo*, 17(2), 228–239.
- Prahartiwi, L. I. (2022). Implementasi Algoritma FP-Growth untuk Menemukan Pola Pembelian Konsumen pada Analisis Keranjang Pasar. *Indonesian Journal on Information System*, 7(1).
- Passolving, I. Y., Ali, P. R., Kartawidjaja, M. A., & Sukwadi, R. (2025). Segmentasi Pelanggan Menggunakan K-Means Clustering: Menganalisis Metrik RFM untuk Strategi Pemasaran. *Jurnal Media Teknik dan Sistem Industri*, 9(1).
- Santoso, E., & Hamdani, A. U. (2023). Segmentasi Pelanggan Pada PT. TAB Menggunakan Algoritma K-Means Clustering. *JIKI (Jurnal Ilmu Komputer dan Informatika)*, 4(2).