

Pemanfaatan Data Mining dengan Metode K-Means Berbasis RFM untuk Segmentasi Pelanggan dalam Mendukung Strategi Customer Intelligence (Studi Kasus: Olist E-Commerce)

Sri Rizki Umanti¹, Muhammad Rafly Setiawan², Ibrahim Musyaffa Rizqie³

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

¹sririzkiunanti@gmail.com, ²mraflisetiawan2205@gmail.com, ³musyaffarizqie2205@gmail.com

Abstract

The rapid growth of the e-commerce sector has generated a substantial volume of customer transaction data that can be utilized as a valuable source of business intelligence. This study aims to apply data mining techniques through the K-Means clustering algorithm combined with the Recency, Frequency, and Monetary (RFM) approach to perform customer segmentation in Olist E-Commerce. The RFM model is employed to evaluate customer behavior based on the most recent transaction date, transaction frequency, and total purchasing value. Furthermore, the K-Means algorithm is utilized to classify customers into groups with similar transaction characteristics. The findings indicate that customer segmentation successfully produces several distinct customer groups, including loyal customers, potential customers, passive customers, and low-value customers. The information generated from the segmentation process can be used to support more targeted and data-driven marketing strategies. In addition, the implementation of the RFM and K-Means methods enables companies to gain a deeper understanding of customer characteristics, thereby enhancing the effectiveness of business strategies and customer relationship management.

Keywords: Data Mining, K-Means, RFM, Customer Segmentation, Customer Intelligence, E-Commerce.

Abstrak

Pertumbuhan industri perdagangan elektronik menghasilkan volume data transaksi pelanggan yang semakin besar dan berpotensi dimanfaatkan sebagai sumber informasi dalam mendukung pengambilan keputusan bisnis. Penelitian ini bertujuan menerapkan teknik data mining melalui algoritma K-Means yang dipadukan dengan pendekatan Recency, Frequency, dan Monetary (RFM) untuk melakukan pengelompokan pelanggan pada platform Olist E-Commerce. Pendekatan RFM digunakan untuk mengevaluasi aktivitas pelanggan berdasarkan waktu transaksi terakhir, frekuensi pembelian, serta nilai transaksi yang dihasilkan. Selanjutnya, algoritma K-Means digunakan untuk mengelompokkan pelanggan yang memiliki karakteristik transaksi serupa ke dalam beberapa segmen. Hasil analisis menghasilkan beberapa kelompok pelanggan dengan pola perilaku yang berbeda, meliputi pelanggan loyal, pelanggan potensial, pelanggan pasif, dan pelanggan bernilai rendah. Informasi yang diperoleh dari proses segmentasi dapat dimanfaatkan sebagai dasar dalam penyusunan strategi pemasaran yang lebih tepat sasaran dan berbasis data. Selain itu, hasil penelitian juga menunjukkan bahwa penerapan metode RFM dan K-Means mampu membantu perusahaan dalam memahami karakteristik pelanggan secara lebih mendalam sehingga mendukung peningkatan efektivitas strategi bisnis dan pengelolaan hubungan pelanggan.

Kata kunci: Data Mining, K-Means, RFM, Segmentasi Pelanggan, Customer Intelligence.

A. PENDAHULUAN

Kemajuan teknologi informasi telah memberikan dampak yang signifikan terhadap berbagai aspek kehidupan dan aktivitas organisasi pertumbuhan industri e-commerce.

Aktivitas transaksi digital yang terus meningkat menghasilkan data dalam jumlah besar yang dapat dimanfaatkan sebagai sumber informasi strategis bagi perusahaan. Data transaksi pelanggan tidak hanya berperan sebagai media penyimpanan data penjualan, tetapi juga dapat digunakan untuk memahami perilaku konsumen dalam melakukan pembelian. Oleh karena itu, perusahaan perlu memanfaatkan data tersebut secara optimal guna meningkatkan efektivitas strategi bisnis dan pemasaran.

Olist E-Commerce merupakan salah satu platform perdagangan digital yang memiliki volume transaksi pelanggan cukup tinggi. Banyaknya data transaksi yang tersimpan dapat diolah untuk mengetahui pola perilaku pelanggan sehingga perusahaan mampu memahami karakteristik konsumennya secara lebih mendalam. Salah satu pendekatan yang dapat diterapkan adalah segmentasi pelanggan menggunakan metode data mining.

Perkembangan teknologi digital telah mengubah cara perusahaan mengelola informasi yang berasal dari aktivitas pelanggan. Setiap transaksi yang dilakukan melalui platform e-commerce menghasilkan data yang terus bertambah dari waktu ke waktu. Apabila data tersebut hanya disimpan sebagai arsip, perusahaan akan kehilangan peluang untuk memperoleh informasi yang bernilai. Oleh karena itu, diperlukan proses analisis yang mampu mengubah kumpulan data transaksi menjadi informasi yang dapat dimanfaatkan sebagai dasar dalam penyusunan strategi bisnis.

Di tengah persaingan bisnis digital yang semakin kompetitif, pemahaman terhadap karakteristik pelanggan menjadi salah satu faktor yang menentukan keberhasilan perusahaan dalam mempertahankan daya saing. Setiap pelanggan memiliki pola transaksi yang berbeda, baik dari sisi waktu pembelian, intensitas transaksi, maupun nilai pembelian yang dilakukan. Perbedaan tersebut menunjukkan bahwa perusahaan tidak dapat menerapkan strategi pemasaran yang sama kepada seluruh pelanggan. Melalui proses segmentasi, perusahaan dapat mengenali kelompok pelanggan yang memiliki karakteristik serupa sehingga program pemasaran dapat disusun secara lebih terarah.

Salah satu teknik yang banyak digunakan untuk melakukan segmentasi pelanggan adalah metode clustering. Teknik ini memungkinkan data dikelompokkan berdasarkan tingkat kemiripan karakteristik tanpa memerlukan label data sebelumnya. Dalam penelitian ini, algoritma K-Means dipilih karena mampu melakukan proses pengelompokan secara efisien pada data transaksi dalam jumlah besar. Agar hasil segmentasi lebih menggambarkan perilaku

pelanggan, proses clustering dikombinasikan dengan pendekatan Recency, Frequency, dan Monetary (RFM). Kombinasi kedua metode tersebut diharapkan mampu menghasilkan informasi yang lebih relevan sebagai pendukung pengambilan keputusan bisnis.

Data mining merupakan proses pengolahan data untuk menemukan pola, hubungan, maupun informasi tersembunyi dari sekumpulan data. Dalam penelitian ini, teknik clustering digunakan untuk mengelompokkan pelanggan berdasarkan tingkat kemiripan perilaku transaksi. Metode clustering yang digunakan adalah K-Metode K-Means digunakan karena dapat melakukan pengelompokan data secara efisien berdasarkan karakteristik tertentu yang dimiliki setiap data.

Penelitian ini juga menggunakan pendekatan Recency, Frequency, Monetary (RFM) sebagai dasar pengukuran perilaku pelanggan. Recency menunjukkan waktu transaksi terakhir pelanggan, sedangkan Frequency menggambarkan jumlah transaksi yang dilakukan, sedangkan Monetary menunjukkan total nilai pembelian pelanggan. Kombinasi metode RFM dan K-Means diharapkan mampu menghasilkan segmentasi pelanggan yang lebih akurat sehingga dapat mendukung strategi customer intelligence dalam perusahaan.

Berdasarkan uraian latar belakang tersebut, penelitian ini dilakukan dengan tujuan mengimplementasikan metode K-Means berbasis RFM dalam melakukan segmentasi pelanggan pada Olist E-Commerce guna mendukung strategi customer intelligence dan meningkatkan efektivitas pengambilan keputusan bisnis berbasis data.

B. METODE

Penelitian ini menerapkan pendekatan kuantitatif dengan memanfaatkan data numerik sebagai dasar analisis memanfaatkan teknik data mining untuk melakukan segmentasi pelanggan berdasarkan data transaksi pada Olist E-Commerce. Metode yang digunakan adalah algoritma K-Means dengan pendekatan Recency, Frequency, Monetary (RFM) untuk mengelompokkan pelanggan berdasarkan pola perilaku transaksi.

Data yang digunakan pada penelitian ini berupa data transaksi pelanggan yang mencakup informasi tanggal transaksi, jumlah transaksi, dan total nilai pembelian pelanggan. Pengumpulan data dilakukan melalui metode dokumentasi dengan memanfaatkan dataset transaksi yang tersedia. Selanjutnya, dilakukan tahap preprocessing data yang meliputi pembersihan data, penghapusan data ganda, penanganan nilai yang kosong, serta normalisasi data agar proses analisis dapat dilakukan secara optimal.

Tahapan penelitian diawali dengan proses identifikasi data transaksi yang akan digunakan sebagai objek penelitian. Seluruh data kemudian melalui proses pemeriksaan untuk memastikan bahwa setiap atribut dapat digunakan pada tahap analisis berikutnya. Data yang tidak lengkap, memiliki nilai ganda, maupun mengandung ketidaksesuaian format terlebih dahulu diperbaiki agar kualitas data tetap terjaga selama proses pengolahan berlangsung.

Setelah proses pembersihan data selesai dilakukan, setiap pelanggan dihitung nilai Recency, Frequency, dan Monetary sebagai representasi perilaku transaksinya. Nilai Recency menunjukkan selang waktu sejak pelanggan melakukan transaksi terakhir. Nilai Frequency menunjukkan banyaknya transaksi yang dilakukan selama periode pengamatan, sedangkan Monetary menunjukkan total nominal pembelian yang telah dilakukan pelanggan. Ketiga atribut tersebut digunakan sebagai dasar dalam proses pembentukan kelompok pelanggan.

Sebelum dilakukan proses clustering, data terlebih dahulu dinormalisasi sehingga seluruh atribut memiliki rentang nilai yang seimbang. Langkah ini bertujuan untuk menghindari dominasi salah satu variabel terhadap proses perhitungan jarak antar data. Selanjutnya, jumlah cluster ditentukan menggunakan metode Elbow dengan memperhatikan perubahan nilai inertia pada beberapa alternatif jumlah cluster. Nilai yang membentuk titik siku dipilih sebagai jumlah cluster yang paling sesuai karena mampu memberikan keseimbangan antara kualitas pengelompokan dan kompleksitas model.

Setelah proses preprocessing selesai, dilakukan perhitungan nilai RFM pada setiap pelanggan. Nilai Recency dihitung berdasarkan jarak waktu sejak transaksi transaksi terakhir pelanggan, sedangkan Frequency diukur berdasarkan jumlah transaksi yang dilakukan oleh pelanggan, sedangkan Monetary dihitung berdasarkan total nominal pembelian pelanggan. Konsep dasar RFM dapat dituliskan sebagai berikut:

$$RFM = (Recency, Frequency, Monetary)$$

Tahap berikutnya adalah proses clustering menggunakan algoritma K-Means. Algoritma ini bekerja dengan mengelompokkan data pelanggan berdasarkan kedekatan karakteristik nilai RFM. Perhitungan jarak antar data dilakukan menggunakan metode Euclidean Distance.

Perhitungan jarak Euclidean (Euclidean Distance) dilakukan menggunakan rumus berikut:

$$d(x, y) = \sum_i = 1n(x_i - y_i)^2$$

Hasil clustering kemudian dianalisis untuk mengetahui karakteristik masing-masing kelompok pelanggan serta menentukan strategi customer intelligence yang sesuai berdasarkan hasil segmentasi yang diperoleh.

C. HASIL DAN PEMBAHASAN

Berdasarkan hasil analisis data yang telah dilakukan transaksi pelanggan Olist E-Commerce menggunakan metode K-Means yang dikombinasikan dengan pendekatan RFM (Recency, Frequency, Monetary), diperoleh segmentasi pelanggan yang terbagi menjadi tiga kelompok berdasarkan karakteristik perilaku transaksi. Penentuan jumlah cluster dilakukan dengan menggunakan metode Elbow Method untuk mengetahui jumlah cluster yang paling optimal berdasarkan nilai inertia. Hasil pengujian menunjukkan bahwa penurunan inertia paling optimal terjadi pada cluster ke-3 sehingga Jumlah cluster yang diterapkan dalam penelitian ini adalah sebanyak tiga cluster.

Hasil pengelompokan data menunjukkan bahwa setiap kelompok pelanggan memiliki karakteristik yang berbeda. Cluster pertama didominasi oleh pelanggan dengan nilai recency tinggi serta frekuensi transaksi dan nilai pembelian yang rendah. Karakteristik tersebut mengindikasikan bahwa pelanggan sudah cukup lama tidak melakukan transaksi sehingga dapat dikategorikan sebagai pelanggan pasif. Cluster kedua terdiri dari pelanggan dengan tingkat transaksi sedang dan memiliki potensi untuk berkembang menjadi pelanggan loyal apabila diberikan strategi pemasaran yang tepat. Sementara itu, cluster ketiga menunjukkan pelanggan dengan frequency dan monetary tinggi serta recency rendah. Kelompok ini merupakan pelanggan loyal yang memberikan kontribusi terbesar terhadap perusahaan.

Analisis hasil clustering menunjukkan bahwa setiap kelompok pelanggan memiliki perilaku transaksi yang berbeda sehingga memerlukan pendekatan pengelolaan yang tidak sama. Perbedaan tersebut terlihat dari kombinasi nilai Recency, Frequency, dan Monetary yang dimiliki oleh masing-masing kelompok. Semakin rendah nilai Recency disertai tingginya nilai Frequency dan Monetary, maka semakin besar kemungkinan pelanggan tersebut termasuk dalam kelompok yang aktif dan memberikan kontribusi tinggi terhadap aktivitas penjualan perusahaan.

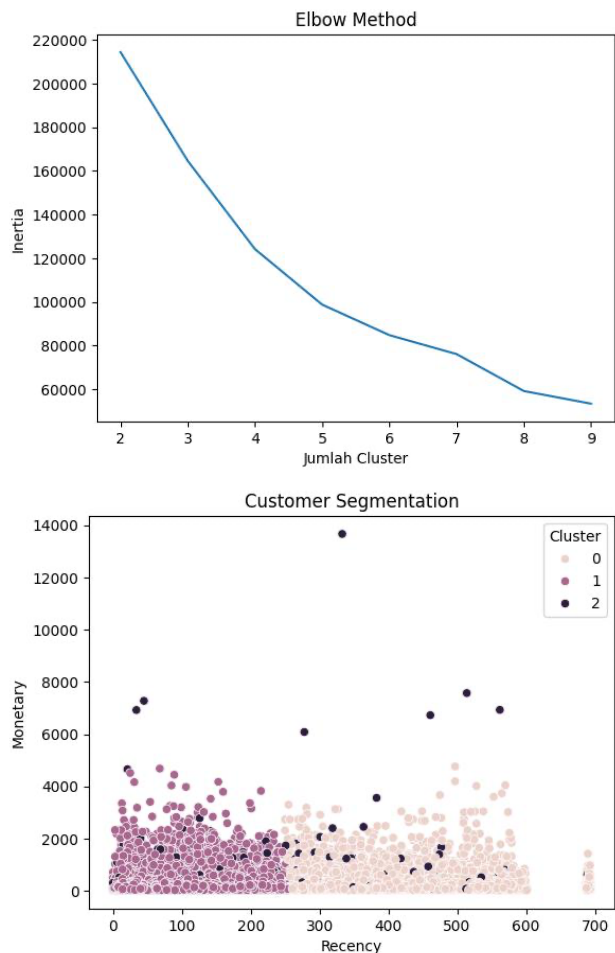
Sebaliknya, pelanggan dengan nilai Recency yang relatif tinggi namun memiliki frekuensi transaksi rendah menunjukkan kecenderungan berkurangnya aktivitas pembelian. Kondisi tersebut dapat menjadi indikator bahwa

pelanggan mulai tidak aktif sehingga perusahaan perlu melakukan strategi retensi, misalnya melalui pemberian promosi khusus, potongan harga, atau penawaran produk yang sesuai dengan riwayat pembelian sebelumnya. Dengan demikian, hasil segmentasi tidak hanya memberikan informasi mengenai karakteristik pelanggan, tetapi juga menjadi dasar dalam penyusunan strategi pemasaran yang lebih efektif.

Pengelompokan pelanggan menggunakan kombinasi metode RFM dan K-Means juga memberikan kemudahan dalam proses analisis data transaksi yang berjumlah besar. Data yang sebelumnya sulit diinterpretasikan dapat disederhanakan menjadi beberapa kelompok dengan karakteristik yang jelas. Informasi tersebut membantu perusahaan dalam menentukan prioritas pelayanan kepada pelanggan, menyusun program loyalitas, serta mengalokasikan anggaran promosi secara lebih efisien sesuai dengan nilai yang diberikan oleh setiap kelompok pelanggan.

Selain itu, hasil segmentasi menunjukkan bahwa pendekatan berbasis data mampu menghasilkan informasi yang lebih objektif dibandingkan keputusan yang hanya didasarkan pada pengalaman atau perkiraan semata. Dengan memanfaatkan hasil clustering, perusahaan dapat lebih mudah mengidentifikasi pelanggan yang memiliki potensi untuk dipertahankan maupun pelanggan yang memerlukan perhatian khusus. Penerapan strategi yang sesuai pada setiap kelompok diharapkan mampu meningkatkan kepuasan pelanggan sekaligus memberikan dampak positif terhadap peningkatan kinerja bisnis secara berkelanjutan.

Hasil segmentasi pelanggan menggunakan metode K-Means berbasis RFM menjadi hasil pengelompokan pelanggan menggunakan metode K-Means berbasis RFM menunjukkan bahwa teknik data mining mampu membantu perusahaan dalam memahami perilaku pelanggan secara lebih mendalam. Informasi hasil clustering dapat dimanfaatkan perusahaan untuk menentukan strategi pemasaran yang lebih optimal, meningkatkan tingkat loyalitas pelanggan, serta mendukung pengambilan keputusan berbasis data. Selain itu, perusahaan juga dapat mengoptimalkan biaya promosi dengan menyesuaikan strategi pemasaran berdasarkan karakteristik masing-masing kelompok pelanggan.



Penerapan metode RFM dan K-Means pada Olist E-Commerce membuktikan bahwa data transaksi pelanggan dapat diolah menjadi informasi strategis yang bermanfaat bagi perusahaan. Informasi tersebut dapat digunakan untuk memahami pola perilaku konsumen sekaligus meningkatkan efektivitas strategi bisnis dan pemasaran perusahaan.

D. Kesimpulan

Berdasarkan hasil penelitian yang telah dilaksanakan, metode K-Means berbasis RFM mampu digunakan secara efektif untuk melakukan segmentasi pelanggan pada Olist E-Commerce. Hasil segmentasi berhasil mengelompokkan pelanggan berdasarkan karakteristik perilaku transaksi sehingga perusahaan dapat memahami pola aktivitas pelanggan secara lebih mendalam.

Penerapan metode ini membantu perusahaan dalam mendukung strategi customer intelligence melalui pemanfaatan data transaksi pelanggan sebagai dasar pengambilan keputusan bisnis. Selain itu, hasil segmentasi juga dapat dimanfaatkan untuk meningkatkan efektivitas pemasaran, mempertahankan loyalitas pelanggan, serta mengoptimalkan strategi promosi berdasarkan karakteristik masing-masing kelompok pelanggan.

Dengan demikian, penerapan teknik data mining menggunakan metode K-Means berbasis RFM mampu memberikan kontribusi positif dalam mendukung pengelolaan hubungan pelanggan dan peningkatan kualitas strategi bisnis pada Olist E-Commerce.

Hasil penelitian ini menunjukkan bahwa penerapan metode K-Means yang dipadukan dengan pendekatan Recency, Frequency, dan Monetary (RFM) mampu menghasilkan segmentasi pelanggan yang dapat dimanfaatkan sebagai pendukung pengambilan keputusan bisnis. Informasi yang dihasilkan dari proses pengelompokan memberikan gambaran mengenai karakteristik masing-masing kelompok pelanggan sehingga perusahaan dapat menyusun strategi pemasaran yang lebih tepat sesuai dengan perilaku transaksi setiap pelanggan.

Penelitian ini masih memiliki ruang untuk dikembangkan pada penelitian selanjutnya, misalnya dengan memanfaatkan jumlah data yang lebih besar, menambahkan atribut lain yang berkaitan dengan perilaku pelanggan, atau membandingkan hasil segmentasi menggunakan algoritma clustering yang berbeda. Pengembangan tersebut diharapkan mampu menghasilkan kualitas segmentasi yang lebih baik sehingga manfaat yang diperoleh perusahaan dalam mendukung strategi customer intelligence dapat semakin optimal.

E. DAFTAR PUSTAKA

- [1] Han, J., Kamber, M., & Pei, J. 2012. Data Mining: Concepts and Techniques. 3rd Edition. Morgan Kaufmann.
- [2] Tan, P. N., Steinbach, M., Karpadne, A., & Kumar, V. 2019. Introduction to Data Mining. 2nd Edition. Pearson Education.
- [3] Larose, D. T., & Larose, C. D. 2014. Discovering Knowledge in Data: An Introduction to Data Mining. 2nd Edition. Wiley.
- [4] Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. 2016. Data Mining: Practical Machine Learning Tools and Techniques. 4th Edition. Morgan Kaufmann.
- [5] Berry, M. J. A., & Linoff, G. S. 2011. Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management. Wiley Publishing.
- [6] Sugiyono. 2019. Metode Penelitian Kuantitatif, Kualitatif, dan R&D. Alfabeta. Bandung.
- [7] Kotler, P., & Keller, K. L. 2016. Marketing Management. 15th Edition. Pearson Education.
- [8] Kotler, P., & Armstrong, G. 2018. Principles of Marketing. 17th Edition. Pearson Education.
- [9] Kumar, V., & Reinartz, W. 2018. Customer Relationship Management: Concept, Strategy, and Tools. Springer.
- [10] Buttle, F., & Maklan, S. 2019. Customer Relationship Management: Concepts and Technologies. 4th Edition. Routledge.