Journal of Data Science and Scientific Modeling
Vol. 1, No. 1, June 2026, pp. 39-47
eISSN 3163-6829 | https://journal.gemacendekia.org/index.php/jdssm

Implementasi Data Mining dalam Identifikasi Pola Perilaku Konsumen pada Platform E-Commerce

Muhammad Farhan Setiawan [a,1,*], Dinda Lestari [a,2]

[a] Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Digital Indonesia, Indonesia
[1] farhansetiawan@gmail.com*; [2] dindalestari@gmail.com
* Corresponding Author

ABSTRACT

The rapid growth of digital technology has significantly accelerated the expansion of the e-commerce sector, generating large and continuously increasing volumes of customer transaction data. This data contains valuable insights that can be utilized to better understand consumer behavior. Data mining is a widely used approach to extract such insights by identifying patterns, trends, and customer characteristics based on transaction activities. This study aims to implement data mining techniques to identify consumer behavior patterns on e-commerce platforms in order to support more effective and targeted marketing strategies. The research method involves collecting customer transaction data, including purchase frequency, transaction value, product categories, transaction time, and customer interaction levels. The data is then processed through cleaning, transformation, and normalization to ensure quality. K-Means clustering is applied to group customers based on similar behavioral characteristics, while the Apriori algorithm is used to analyze associations between products frequently purchased together. The results identify distinct customer segments, including loyal customers, moderate buyers with specific preferences, and sporadic low-value buyers. Association analysis also reveals product bundling patterns, supporting cross-selling and recommendation strategies, and improving overall marketing effectiveness.

Copyright © 2026, The Author(s)
This is an open-access article under the CC–BY-SA license

Article History
Received 2026-03-26
Revised 2026-04-28
Accepted 2026-05-05
Published 2026-06-24

Keywords
Data Mining;
E Commerce;
Perilaku Konsumen;
K-Means Clustering;
Algoritma Apriori

PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi dalam dua dekade terakhir telah membawa perubahan signifikan terhadap berbagai aspek kehidupan masyarakat, termasuk dalam aktivitas perdagangan dan bisnis. Transformasi digital yang terjadi secara masif telah mendorong munculnya berbagai platform perdagangan elektronik atau electronic commerce (e-commerce) yang memungkinkan transaksi dilakukan secara cepat, mudah, dan tanpa batasan geografis. Keberadaan platform e-commerce tidak hanya mengubah pola konsumsi masyarakat, tetapi juga menghasilkan data dalam jumlah besar yang berasal dari aktivitas pelanggan selama proses pencarian, pemilihan, hingga pembelian produk. Data tersebut menjadi aset strategis yang dapat dimanfaatkan oleh perusahaan untuk memahami perilaku konsumen secara lebih mendalam (Han et al., 2022).

Pertumbuhan industri e-commerce di Indonesia menunjukkan tren yang sangat positif dalam beberapa tahun terakhir. Meningkatnya penetrasi internet, penggunaan perangkat seluler, serta perubahan gaya hidup masyarakat yang semakin mengandalkan layanan digital telah mendorong peningkatan jumlah transaksi online secara signifikan. Kondisi ini menyebabkan perusahaan e-commerce harus mampu mengelola dan menganalisis data pelanggan secara efektif agar dapat mempertahankan daya saing di tengah persaingan pasar yang semakin ketat. Data pelanggan yang tersimpan dalam sistem perusahaan mengandung informasi penting yang dapat

How to cite: Setiawan, M. F., & Lestari, D. (2026). Implementasi Data Mining dalam Identifikasi Pola Perilaku Konsumen pada Platform E-Commerce. *Journal of Data Science and Scientific Modeling, 1*(1), 39-47. https://doi.org/10.71094/jdssm.v1i1.77

digunakan untuk mengidentifikasi kebutuhan, preferensi, serta kebiasaan pembelian konsumen (Laudon & Traver, 2021).

Dalam lingkungan bisnis digital yang sangat kompetitif, pemahaman terhadap perilaku konsumen menjadi faktor penting dalam menentukan keberhasilan strategi pemasaran perusahaan. Perusahaan tidak lagi hanya mengandalkan intuisi atau pengalaman dalam mengambil keputusan bisnis, melainkan memanfaatkan pendekatan berbasis data untuk memperoleh wawasan yang lebih akurat dan objektif. Analisis perilaku konsumen memungkinkan perusahaan untuk mengetahui karakteristik pelanggan, pola pembelian, tingkat loyalitas, serta kecenderungan konsumen dalam memilih produk atau layanan tertentu. Informasi tersebut dapat digunakan untuk meningkatkan kualitas layanan dan efektivitas strategi pemasaran yang diterapkan (Kotler et al., 2022).

Seiring dengan meningkatnya volume data yang dihasilkan oleh aktivitas transaksi digital, diperlukan metode yang mampu mengekstraksi informasi penting dari kumpulan data yang besar dan kompleks. Salah satu pendekatan yang banyak digunakan dalam bidang analisis data adalah data mining. Data mining merupakan proses menemukan pola, hubungan, dan pengetahuan yang tersembunyi dalam kumpulan data berukuran besar melalui penerapan berbagai teknik statistik, matematika, dan kecerdasan buatan. Penerapan data mining telah terbukti mampu membantu organisasi dalam mengidentifikasi tren bisnis, memprediksi perilaku pelanggan, serta mendukung proses pengambilan keputusan yang lebih efektif (Witten et al., 2017; Zaki & Meira, 2020).

Data mining memiliki berbagai teknik yang dapat digunakan sesuai dengan tujuan analisis yang ingin dicapai. Salah satu teknik yang sering digunakan dalam analisis perilaku konsumen adalah clustering. Clustering merupakan metode pengelompokan data berdasarkan tingkat kemiripan karakteristik tertentu sehingga objek yang berada dalam satu kelompok memiliki kesamaan yang lebih tinggi dibandingkan dengan objek pada kelompok lain. Teknik ini memungkinkan perusahaan untuk melakukan segmentasi pelanggan secara otomatis berdasarkan data transaksi yang tersedia. Segmentasi pelanggan menjadi langkah penting dalam pengembangan strategi pemasaran yang lebih personal dan tepat sasaran (Tan et al., 2019).

Algoritma K-Means merupakan salah satu metode clustering yang paling populer karena memiliki tingkat efisiensi yang tinggi serta mudah diimplementasikan pada berbagai jenis data. Algoritma ini bekerja dengan mengelompokkan data ke dalam sejumlah cluster berdasarkan jarak terdekat terhadap pusat cluster (centroid). Dalam konteks e-commerce, K-Means dapat digunakan untuk mengidentifikasi kelompok pelanggan berdasarkan frekuensi pembelian, nilai transaksi, maupun preferensi produk yang dibeli. Hasil segmentasi tersebut dapat menjadi dasar bagi perusahaan dalam menyusun program promosi yang lebih efektif dan sesuai dengan karakteristik masing-masing kelompok pelanggan (Jain, 2010).

Selain clustering, analisis asosiasi juga menjadi salah satu teknik data mining yang banyak diterapkan dalam industri e-commerce. Analisis asosiasi bertujuan untuk menemukan hubungan antaritem dalam suatu transaksi sehingga dapat diketahui produk-produk yang sering dibeli secara bersamaan oleh pelanggan. Informasi tersebut sangat berguna dalam mendukung strategi pemasaran seperti cross-selling, up-selling, serta pengembangan sistem rekomendasi produk yang lebih relevan bagi pelanggan. Salah satu algoritma yang umum digunakan dalam analisis asosiasi adalah algoritma Apriori yang mampu menghasilkan aturan asosiasi berdasarkan nilai support dan confidence tertentu (Agrawal et al., 1993).

Pemanfaatan algoritma Apriori dalam platform e-commerce memberikan keuntungan yang signifikan bagi perusahaan dalam memahami pola pembelian konsumen. Dengan mengetahui kombinasi produk yang sering dibeli secara bersamaan, perusahaan dapat menyusun tata letak

produk yang lebih efektif, menawarkan paket promosi yang menarik, serta meningkatkan peluang terjadinya pembelian tambahan. Selain itu, hasil analisis asosiasi juga dapat digunakan untuk mengembangkan sistem rekomendasi yang mampu meningkatkan pengalaman pelanggan selama berbelanja secara online (Maimon & Rokach, 2020).

Berbagai penelitian sebelumnya menunjukkan bahwa penerapan data mining pada sektor e-commerce mampu memberikan kontribusi positif terhadap peningkatan kinerja bisnis. Penelitian yang dilakukan oleh Kotu dan Deshpande (2019) menunjukkan bahwa teknik clustering dapat digunakan untuk mengidentifikasi segmen pelanggan dengan karakteristik yang berbeda sehingga perusahaan dapat menerapkan strategi pemasaran yang lebih efektif. Sementara itu, penelitian lain menunjukkan bahwa analisis asosiasi mampu meningkatkan akurasi rekomendasi produk serta mendorong peningkatan nilai transaksi pelanggan melalui strategi pemasaran yang lebih terarah (Han et al., 2022).

Meskipun berbagai penelitian telah membahas penerapan data mining dalam sektor perdagangan elektronik, masih terdapat kebutuhan untuk mengeksplorasi lebih lanjut bagaimana kombinasi antara teknik clustering dan analisis asosiasi dapat digunakan secara terpadu dalam memahami perilaku konsumen. Integrasi kedua metode tersebut diharapkan mampu menghasilkan informasi yang lebih komprehensif mengenai karakteristik pelanggan dan pola pembelian yang terjadi pada platform e-commerce. Dengan demikian, perusahaan dapat memperoleh wawasan yang lebih mendalam dalam menyusun strategi bisnis berbasis data.

Penelitian ini dilaksanakan dengan memanfaatkan data transaksi pelanggan yang mencakup frekuensi pembelian, nilai transaksi, kategori produk, serta riwayat aktivitas pelanggan pada platform e-commerce. Teknik clustering menggunakan algoritma K-Means digunakan untuk mengelompokkan pelanggan berdasarkan karakteristik perilaku pembelian, sedangkan algoritma Apriori diterapkan untuk mengidentifikasi hubungan antarproduk yang sering muncul dalam satu transaksi. Kombinasi kedua metode tersebut diharapkan mampu memberikan gambaran yang lebih jelas mengenai pola perilaku konsumen pada lingkungan bisnis digital.

Urgensi penelitian ini terletak pada pentingnya pemanfaatan data sebagai dasar pengambilan keputusan dalam industri e-commerce yang terus berkembang. Di tengah persaingan bisnis yang semakin kompetitif, perusahaan dituntut untuk mampu memahami kebutuhan dan perilaku pelanggan secara lebih akurat agar dapat memberikan layanan yang lebih personal dan meningkatkan loyalitas pelanggan. Oleh karena itu, penerapan data mining menjadi salah satu pendekatan yang relevan untuk mendukung transformasi bisnis berbasis data dan meningkatkan keunggulan kompetitif perusahaan.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk mengimplementasikan teknik data mining melalui algoritma K-Means dan Apriori dalam mengidentifikasi pola perilaku konsumen pada platform e-commerce. Hasil penelitian diharapkan dapat memberikan kontribusi teoritis dalam pengembangan kajian data mining pada sektor perdagangan elektronik serta memberikan manfaat praktis bagi perusahaan dalam merancang strategi pemasaran yang lebih efektif, efisien, dan berbasis data.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis data mining untuk mengidentifikasi pola perilaku konsumen pada platform e-commerce. Pendekatan kuantitatif dipilih karena mampu memberikan gambaran yang objektif mengenai karakteristik dan kecenderungan perilaku pelanggan berdasarkan data transaksi yang tersedia. Metode ini

memungkinkan peneliti melakukan pengolahan data dalam jumlah besar secara sistematis untuk menghasilkan informasi yang dapat digunakan sebagai dasar pengambilan keputusan bisnis (Han et al., 2022; Kotu & Deshpande, 2019).

Desain penelitian yang digunakan adalah penelitian deskriptif dan eksploratif. Penelitian deskriptif bertujuan untuk menggambarkan karakteristik perilaku konsumen berdasarkan data transaksi yang tersedia, sedangkan penelitian eksploratif digunakan untuk menemukan pola-pola tersembunyi yang belum diketahui sebelumnya melalui penerapan teknik data mining. Kombinasi kedua pendekatan tersebut dianggap sesuai untuk menghasilkan pemahaman yang lebih komprehensif mengenai perilaku pelanggan pada lingkungan e-commerce yang dinamis (Larose & Larose, 2015; Provost & Fawcett, 2013).

Objek penelitian adalah data transaksi pelanggan pada platform e-commerce yang mencakup informasi frekuensi pembelian, nilai transaksi, kategori produk yang dibeli, waktu pembelian, metode pembayaran, dan aktivitas pelanggan selama periode penelitian. Data yang digunakan merupakan data sekunder yang diperoleh dari basis data transaksi perusahaan atau sumber data simulasi yang memiliki karakteristik serupa dengan data riil. Penggunaan data transaksi memungkinkan analisis perilaku pelanggan dilakukan secara faktual berdasarkan aktivitas aktual yang tercatat dalam sistem (Tan et al., 2019).

Tahap pertama penelitian adalah pengumpulan data (data collection). Data transaksi pelanggan dikumpulkan dari sistem basis data e-commerce menggunakan teknik ekstraksi data. Pada tahap ini dilakukan identifikasi atribut yang relevan dengan tujuan penelitian, seperti identitas pelanggan, tanggal transaksi, jumlah produk yang dibeli, total nilai pembelian, kategori produk, serta riwayat interaksi pelanggan. Pengumpulan data yang terstruktur menjadi faktor penting dalam menghasilkan analisis yang akurat dan dapat dipercaya (Witten et al., 2017; Shmueli et al., 2020).

Setelah data terkumpul, dilakukan proses pembersihan data (data cleaning). Tahap ini bertujuan untuk menghilangkan data yang tidak lengkap, duplikasi data, kesalahan pencatatan, serta nilai yang tidak konsisten. Data yang memiliki nilai kosong (missing values) ditangani menggunakan metode imputasi atau eliminasi sesuai tingkat kekurangan data yang ditemukan. Pembersihan data merupakan langkah penting karena kualitas hasil analisis sangat bergantung pada kualitas data yang digunakan (García et al., 2016; Kelleher et al., 2020).

Tahap berikutnya adalah transformasi dan normalisasi data. Data transaksi yang memiliki skala berbeda diubah ke dalam format yang seragam menggunakan teknik normalisasi Min-Max atau Z-Score. Proses transformasi juga dilakukan terhadap atribut kategorikal agar dapat diolah menggunakan algoritma data mining. Tahapan ini bertujuan untuk meningkatkan kinerja algoritma serta mengurangi bias yang dapat muncul akibat perbedaan skala data (Aggarwal, 2015; James et al., 2021).

Untuk mengidentifikasi kelompok pelanggan yang memiliki karakteristik serupa, penelitian ini menggunakan teknik clustering dengan algoritma K-Means. Algoritma K-Means dipilih karena memiliki kemampuan yang baik dalam mengelompokkan data berdasarkan kedekatan karakteristik antarobjek. Penentuan jumlah cluster dilakukan menggunakan metode Elbow Method yang bertujuan untuk memperoleh jumlah kelompok yang paling optimal. Hasil clustering kemudian dianalisis untuk mengidentifikasi segmen pelanggan berdasarkan perilaku pembelian mereka (Jain, 2010; Han et al., 2022).

Variabel yang digunakan dalam proses clustering mengacu pada pendekatan RFM (Recency, Frequency, Monetary). Variabel recency mengukur jarak waktu sejak transaksi terakhir pelanggan, frequency mengukur jumlah transaksi yang dilakukan pelanggan dalam periode tertentu, sedangkan monetary menunjukkan total nilai transaksi pelanggan. Pendekatan RFM

telah banyak digunakan dalam penelitian perilaku konsumen karena mampu menggambarkan tingkat loyalitas dan nilai pelanggan secara efektif (Kumar & Reinartz, 2018; Buttle & Maklan, 2019).

Selain clustering, penelitian ini juga menerapkan analisis asosiasi menggunakan algoritma Apriori untuk menemukan pola keterkaitan antarproduk yang sering dibeli secara bersamaan oleh pelanggan. Algoritma Apriori bekerja dengan mencari frequent itemsets berdasarkan nilai minimum support dan confidence yang telah ditentukan sebelumnya. Hasil analisis asosiasi digunakan untuk mengidentifikasi peluang strategi cross-selling, bundling produk, serta pengembangan sistem rekomendasi produk pada platform e-commerce (Agrawal & Srikant, 1994; Tan et al., 2019).

Evaluasi hasil clustering dilakukan menggunakan Silhouette Coefficient untuk mengukur kualitas pengelompokan data. Nilai silhouette yang mendekati angka satu menunjukkan bahwa objek dalam suatu cluster memiliki tingkat kesamaan yang tinggi dan berbeda secara signifikan dengan cluster lainnya. Sementara itu, evaluasi aturan asosiasi dilakukan dengan memperhatikan nilai support, confidence, dan lift ratio untuk memastikan bahwa hubungan antarproduk yang ditemukan memiliki tingkat signifikansi yang tinggi (Kotu & Deshpande, 2019; Witten et al., 2017).

Proses pengolahan dan analisis data dilakukan menggunakan perangkat lunak Python dengan dukungan pustaka Pandas, NumPy, Scikit-Learn, dan Mlxtend. Penggunaan perangkat lunak tersebut memungkinkan pengolahan data dalam jumlah besar secara efisien serta mendukung implementasi berbagai algoritma data mining yang dibutuhkan dalam penelitian. Visualisasi hasil analisis dilakukan menggunakan Matplotlib dan Seaborn untuk memudahkan interpretasi pola perilaku konsumen yang ditemukan (McKinney, 2022; Pedregosa et al., 2011; Müller & Guido, 2017).

Tahap akhir penelitian adalah interpretasi hasil dan penyusunan rekomendasi bisnis. Hasil segmentasi pelanggan dan pola asosiasi produk dianalisis untuk menghasilkan rekomendasi strategis yang dapat diterapkan oleh perusahaan e-commerce. Rekomendasi tersebut meliputi strategi personalisasi promosi, pengembangan program loyalitas pelanggan, optimasi penempatan produk, serta peningkatan efektivitas sistem rekomendasi. Dengan demikian, hasil penelitian diharapkan mampu memberikan kontribusi bagi pengembangan strategi pemasaran berbasis data serta meningkatkan daya saing perusahaan di era digital yang semakin kompetitif (Provost & Fawcett, 2013; Kumar & Reinartz, 2018).

HASIL DAN PEMBAHASAN

Hasil penelitian menunjukkan bahwa penerapan teknik data mining pada data transaksi pelanggan e-commerce mampu menghasilkan informasi yang signifikan terkait pola perilaku konsumen. Setelah melalui proses pembersihan data (data cleaning), transformasi data, dan normalisasi, diperoleh dataset yang layak untuk dianalisis menggunakan algoritma K-Means dan Apriori. Dataset yang digunakan terdiri atas berbagai atribut, seperti frekuensi pembelian, nilai transaksi, kategori produk, waktu pembelian, dan tingkat interaksi pelanggan dengan platform. Tahapan pra-pemrosesan data terbukti sangat penting karena mampu mengurangi noise dan meningkatkan kualitas hasil analisis. Temuan ini sejalan dengan penelitian yang menyatakan bahwa kualitas data menjadi faktor utama dalam keberhasilan implementasi data mining pada sistem bisnis digital (Prasetyo & Nugroho, 2021).

Berdasarkan hasil pengelompokan menggunakan algoritma K-Means, diperoleh tiga kelompok pelanggan utama yang memiliki karakteristik berbeda. Kelompok pertama terdiri atas

pelanggan dengan frekuensi transaksi tinggi dan nilai pembelian yang relatif besar. Pelanggan dalam kelompok ini menunjukkan tingkat loyalitas yang tinggi terhadap platform e-commerce. Mereka cenderung melakukan pembelian secara berulang dalam periode tertentu dan memiliki ketertarikan terhadap berbagai kategori produk. Kelompok ini menjadi segmen yang sangat potensial untuk dipertahankan melalui program loyalitas pelanggan dan pemberian insentif khusus.

Kelompok kedua terdiri atas pelanggan dengan frekuensi pembelian sedang dan pola konsumsi yang lebih spesifik terhadap kategori produk tertentu. Pelanggan pada segmen ini umumnya memiliki preferensi yang jelas terhadap produk tertentu dan menunjukkan kecenderungan melakukan pembelian ulang pada kategori yang sama. Karakteristik tersebut menunjukkan bahwa strategi pemasaran yang bersifat personalisasi akan lebih efektif diterapkan pada kelompok ini. Pemanfaatan sistem rekomendasi berbasis riwayat pembelian dapat meningkatkan peluang transaksi serta kepuasan pelanggan.

Sementara itu, kelompok ketiga merupakan pelanggan dengan frekuensi transaksi rendah dan nilai pembelian yang relatif kecil. Segmen ini menunjukkan tingkat keterlibatan yang rendah terhadap platform e-commerce. Hasil analisis menunjukkan bahwa sebagian besar pelanggan pada kelompok ini hanya melakukan transaksi pada saat terdapat promosi atau diskon tertentu. Kondisi tersebut mengindikasikan bahwa faktor harga masih menjadi pertimbangan utama dalam keputusan pembelian mereka. Oleh karena itu, strategi pemasaran berupa promosi berkala dan pemberian voucher diskon dinilai dapat meningkatkan aktivitas transaksi pada kelompok ini.

Analisis lebih lanjut menggunakan algoritma Apriori menghasilkan sejumlah aturan asosiasi (association rules) yang menunjukkan hubungan antarproduk yang sering dibeli secara bersamaan. Salah satu pola yang ditemukan adalah kecenderungan pelanggan membeli produk elektronik bersamaan dengan aksesori pendukungnya. Selain itu, produk kebutuhan rumah tangga juga menunjukkan tingkat asosiasi yang cukup tinggi dengan produk konsumsi harian. Temuan ini memberikan informasi yang sangat berguna bagi perusahaan dalam menyusun strategi cross-selling dan bundling produk untuk meningkatkan nilai transaksi pelanggan.

Nilai support dan confidence yang dihasilkan dari analisis Apriori menunjukkan bahwa beberapa kombinasi produk memiliki tingkat keterkaitan yang kuat. Semakin tinggi nilai confidence yang diperoleh, semakin besar kemungkinan pelanggan membeli produk tertentu setelah membeli produk lainnya. Hasil ini mendukung penerapan sistem rekomendasi otomatis yang dapat ditampilkan kepada pelanggan selama proses transaksi berlangsung. Penelitian sebelumnya menunjukkan bahwa penggunaan teknik association rule mining mampu meningkatkan efektivitas pemasaran digital serta memperbesar peluang pembelian tambahan oleh konsumen (Wibowo et al., 2022).

Dari perspektif bisnis, hasil segmentasi pelanggan memberikan manfaat yang signifikan dalam penyusunan strategi Customer Relationship Management (CRM). Perusahaan dapat merancang pendekatan pemasaran yang berbeda untuk setiap kelompok pelanggan berdasarkan karakteristik yang dimiliki. Segmentasi yang akurat memungkinkan perusahaan mengalokasikan sumber daya pemasaran secara lebih efisien dan meningkatkan tingkat konversi penjualan. Hal ini sesuai dengan pendapat Santoso dan Hidayat (2020) yang menyatakan bahwa segmentasi pelanggan berbasis data merupakan fondasi penting dalam pengembangan strategi pemasaran modern.

Selain membantu strategi pemasaran, penerapan data mining juga berkontribusi dalam proses pengambilan keputusan manajerial. Informasi yang dihasilkan dari analisis data transaksi dapat digunakan untuk memprediksi tren pembelian di masa mendatang. Dengan demikian, perusahaan dapat melakukan perencanaan stok yang lebih baik, mengurangi risiko kelebihan atau

kekurangan persediaan, serta meningkatkan efisiensi operasional. Kemampuan prediktif ini menjadi salah satu keunggulan utama penerapan data science dalam industri e-commerce yang terus berkembang pesat.

Hasil penelitian juga menunjukkan bahwa integrasi data mining dengan sistem informasi e-commerce mampu meningkatkan pemahaman perusahaan terhadap kebutuhan dan preferensi pelanggan. Data yang sebelumnya hanya berfungsi sebagai arsip transaksi dapat diolah menjadi sumber informasi strategis yang mendukung inovasi bisnis. Dengan memanfaatkan teknologi analitik yang tepat, perusahaan dapat merespons perubahan perilaku konsumen secara lebih cepat dan akurat. Temuan ini memperkuat hasil penelitian Rahmawati dan Sari (2023) yang menyatakan bahwa pemanfaatan big data analytics mampu meningkatkan daya saing perusahaan dalam lingkungan bisnis digital.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan. Data yang digunakan masih terbatas pada periode waktu tertentu sehingga belum sepenuhnya merepresentasikan perubahan perilaku pelanggan dalam jangka panjang. Selain itu, algoritma K-Means memiliki sensitivitas terhadap penentuan jumlah cluster yang dapat memengaruhi hasil segmentasi. Oleh karena itu, penelitian selanjutnya dapat mengombinasikan metode clustering lain seperti DBSCAN atau Hierarchical Clustering untuk memperoleh hasil yang lebih optimal (Xu & Wunsch, 2009).

Secara keseluruhan, hasil penelitian membuktikan bahwa implementasi data mining pada platform e-commerce mampu memberikan nilai tambah yang signifikan bagi perusahaan. Teknik clustering berhasil mengidentifikasi segmen pelanggan yang berbeda, sedangkan algoritma Apriori mampu menemukan pola hubungan antarproduk yang bernilai strategis. Informasi yang diperoleh dapat digunakan untuk meningkatkan efektivitas pemasaran, memperkuat loyalitas pelanggan, mengoptimalkan manajemen persediaan, serta mendukung pengambilan keputusan bisnis berbasis data. Dengan semakin meningkatnya volume data transaksi digital, pemanfaatan data mining diperkirakan akan menjadi kebutuhan utama bagi perusahaan e-commerce dalam menghadapi persaingan bisnis yang semakin kompetitif (Putra & Firmansyah, 2021; Turban et al., 2018).

KESIMPULAN

Berdasarkan hasil penelitian, penerapan teknik data mining pada platform e-commerce terbukti mampu mengidentifikasi pola perilaku konsumen secara efektif melalui pengolahan dan analisis data transaksi pelanggan. Pemanfaatan metode clustering dengan algoritma K-Means memungkinkan pengelompokan pelanggan ke dalam beberapa segmen yang memiliki karakteristik pembelian yang berbeda, sehingga perusahaan dapat memahami profil konsumennya secara lebih mendalam. Informasi yang dihasilkan dari proses segmentasi tersebut menjadi dasar penting dalam penyusunan strategi pemasaran yang lebih terarah dan sesuai dengan kebutuhan masing-masing kelompok pelanggan.

Selain itu, penggunaan algoritma Apriori dalam analisis asosiasi berhasil mengungkap hubungan antarproduk yang sering dibeli secara bersamaan oleh pelanggan. Temuan ini memberikan manfaat yang signifikan bagi perusahaan dalam merancang strategi cross-selling, up-selling, serta sistem rekomendasi produk yang lebih relevan. Dengan memahami pola keterkaitan produk, platform e-commerce dapat meningkatkan peluang penjualan sekaligus memberikan pengalaman berbelanja yang lebih baik kepada pelanggan.

Hasil penelitian juga menunjukkan bahwa pelanggan pada platform e-commerce memiliki karakteristik perilaku yang beragam, mulai dari pelanggan loyal dengan frekuensi dan nilai

transaksi tinggi, pelanggan dengan preferensi produk tertentu, hingga pelanggan dengan aktivitas pembelian yang relatif rendah. Perbedaan karakteristik tersebut menegaskan pentingnya pendekatan pemasaran yang berbasis data agar perusahaan dapat memberikan layanan, promosi, dan penawaran yang sesuai dengan kebutuhan setiap segmen pelanggan. Strategi yang tepat akan berkontribusi pada peningkatan kepuasan pelanggan dan penguatan loyalitas jangka panjang.

Secara keseluruhan, implementasi data mining merupakan solusi yang efektif dalam mendukung pengambilan keputusan bisnis pada sektor e-commerce. Kemampuan data mining dalam mengekstraksi informasi yang bernilai dari data transaksi yang besar memungkinkan perusahaan memperoleh wawasan strategis untuk meningkatkan efektivitas pemasaran, mengoptimalkan pengelolaan pelanggan, serta memperkuat daya saing bisnis di era digital. Oleh karena itu, penerapan teknologi data mining perlu terus dikembangkan dan diintegrasikan ke dalam sistem bisnis modern guna mendukung pertumbuhan dan keberlanjutan perusahaan di masa mendatang.

DAFTAR PUSTAKA

Aggarwal, C. C. (2015). *Data mining: The textbook*. Springer.

Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. *Proceedings of the 20th International Conference on Very Large Data Bases*, 487–499.

Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, 207–216.

Buttle, F., & Maklan, S. (2019). *Customer relationship management: Concepts and technologies* (4th ed.). Routledge.

García, S., Luengo, J., & Herrera, F. (2016). *Data preprocessing in data mining*. Springer.

Han, J., Pei, J., & Kamber, M. (2022). *Data mining: Concepts and techniques* (4th ed.). Morgan Kaufmann.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters, 31*(8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (2nd ed.). Springer.

Kelleher, J. D., Tierney, B., & Namee, B. M. (2020). *Data science*. MIT Press.

Kotler, P., Keller, K. L., & Chernev, A. (2022). *Marketing management* (16th ed.). Pearson Education.

Kotu, V., & Deshpande, B. (2019). *Predictive analytics and data mining: Concepts and practice with rapidminer*. Morgan Kaufmann.

Kumar, V., & Reinartz, W. (2018). *Customer relationship management: Concept, strategy, and tools* (3rd ed.). Springer.

Larose, D. T., & Larose, C. D. (2015). *Data mining and predictive analytics* (2nd ed.). Wiley.

Laudon, K. C., & Traver, C. G. (2021). *E-commerce: Business, technology, society* (17th ed.). Pearson.

Maimon, O., & Rokach, L. (2020). *Data mining and knowledge discovery handbook* (3rd ed.). Springer.

McKinney, W. (2022). *Python for data analysis* (3rd ed.). O'Reilly Media.

Müller, A. C., & Guido, S. (2017). *Introduction to machine learning with Python*. O'Reilly Media.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher,

M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Prasetyo, A., & Nugroho, B. (2021). Implementasi data mining untuk analisis perilaku pelanggan pada sistem perdagangan elektronik. *Jurnal Teknologi Informasi dan Komunikasi, 12*(2), 115–126.

Provost, F., & Fawcett, T. (2013). *Data science for business*. O'Reilly Media.

Putra, D. R., & Firmansyah, M. (2021). Pemanfaatan data science dalam meningkatkan kinerja bisnis digital di Indonesia. *Jurnal Sistem Informasi Bisnis, 11*(3), 201–213.

Rahmawati, E., & Sari, N. (2023). Analisis big data terhadap strategi pemasaran pada platform e-commerce. *Jurnal Informatika dan Bisnis Digital, 8*(1), 45–57.

Santoso, H., & Hidayat, R. (2020). Segmentasi pelanggan menggunakan metode clustering untuk mendukung customer relationship management. *Jurnal Teknologi dan Manajemen Informasi, 6*(2), 87–98.

Shmueli, G., Bruce, P. C., Gedeck, P., & Patel, N. R. (2020). *Data mining for business analytics: Concepts, techniques, and applications in Python*. Wiley.

Tan, P. N., Steinbach, M., Karpatne, A., & Kumar, V. (2019). *Introduction to data mining* (2nd ed.). Pearson.

Turban, E., Pollard, C., & Wood, G. (2018). *Information technology for management: On-demand strategies for performance, growth and sustainability*. Wiley.

Wibowo, A., Pramono, Y., & Kurniawan, F. (2022). Penerapan algoritma Apriori dalam analisis pola pembelian konsumen pada marketplace online. *Jurnal Ilmu Komputer Indonesia, 7*(4), 233–245.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). *Data mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann.

Xu, R., & Wunsch, D. (2009). *Clustering*. Wiley-IEEE Press.

Zaki, M. J., & Meira, W. (2020). *Data mining and machine learning: Fundamental concepts and algorithms* (2nd ed.). Cambridge University Press.