Indonesian Research Journal on Education

e-ISSN: 2775 – 8672, p-ISSN: 2775 – 9482
Volume 5 Nomor 4 Tahun 2025
Halaman: 1717 – 1720

Web: https://irje.org/index.php/irje

RESEARCH ARTICLE

Perbandingan K-Means dan K-Medoids dalam Segmentasi Pelanggan E-Commerce: System Literature Review

Mosses Ara'al De Sela[1], Debora Hizkhia Pardede[2], Elkin Rilvani[3]
Universitas Pelita Bangsa, Indonesia

araal2612@gmail.com[1], deborahizkhia@gmail.com[2], elkin.rivalni@pelitabangsa.ac.id[3]

Kata Kunci:
Tinjauan Literatur Sistematis, K-Means, K-Medoids, Pengelompokan, Segmentasi

Abstrak

Segmentasi pelanggan merupakan salah satu strategi penting dalam e-commerce untuk memahami perilaku konsumen dan meningkatkan efektivitas pemasaran. Dua algoritma yang sering digunakan dalam segmentasi adalah K-Means dan K-Medoids. Penelitian ini bertujuan untuk melakukan tinjauan sistematis terhadap literatur yang membandingkan performa K-Means dan K-Medoids dalam konteks segmentasi pelanggan e-commerce. Metode systematic literature review (SLR) digunakan dengan mengacu pada protokol pencarian, seleksi, dan analisis dari berbagai sumber ilmiah terindeks selama 5 tahun terakhir. Hasil tinjauan menunjukkan bahwa K-Means unggul dalam efisiensi komputasi dan cocok untuk dataset besar, namun lebih sensitif terhadap Outlier. Sebaliknya, K-Medoids menunjukkan ketahanan lebih tinggi terhadap Outlier dan hasil klaster yang lebih stabil, meskipun dengan beban komputasi yang lebih besar. Pemilihan algoritma yang tepat bergantung pada karakteristik data dan tujuan segmentasi. Studi ini memberikan wawasan penting bagi peneliti dan praktisi dalam menentukan pendekatan segmentasi pelanggan yang paling sesuai dalam lingkungan e-commerce.

Key Word:
Systematic Literature Review, K-Means,K-Medoids, Clustering, Segmentasi

Abstract

Customer segmentation is a crucial strategy in e-commerce for understanding consumer behavior and enhancing marketing effectiveness. Two commonly used algorithms in segmentation are K-Means and K-Medoids. This study aims to conduct a systematic literature review (SLR) of scholarly works that compare the performance of K-Means and K-Medoids in the context of customer segmentation within e-commerce. The SLR method was carried out based on structured protocols for searching, selecting, and analyzing relevant academic sources indexed within the last five years. The review findings indicate that K-Means excels in computational efficiency and is well-suited for large datasets, but it is more sensitive to Outliers. In contrast, K-Medoids demonstrates greater robustness against Outliers and produces more stable cluster results, albeit with higher computational costs. The selection of the appropriate algorithm depends on data characteristics and segmentation objectives. This study provides valuable insights for researchers and practitioners in selecting the most suitable customer segmentation approach in the dynamic e-commerce environment.

Copyright © 2025 Mosses Ara'al De Sela, Debora Hizkhia Pardede, Elkin Rilvani
*This work is licensed under an **Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)***

PENDAHULUAN

Seiring dengan pesatnya kemajuan teknologi informasi, aktivitas belanja online telah menjadi bagian yang tidak terpisahkan dari gaya hidup masyarakat saat ini. Hal ini terjadi karena berbagai kemudahan yang ditawarkan oleh e-commerce dalam memenuhi kebutuhan konsumen, mulai dari kebutuhan dasar hingga keinginan tambahan. Oleh sebab itu, memahami perilaku konsumen menjadi aspek penting dalam merancang strategi pemasaran yang tepat dan efektif. Dalam hal ini, segmentasi pasar menjadi strategi utama yang berperan dalam mengenali perbedaan kebutuhan serta preferensi konsumen secara lebih terfokus.

Segmentasi pelanggan adalah pendekatan yang bertujuan untuk membagi konsumen ke dalam beberapa kelompok berdasarkan karakteristik, perilaku, atau kebutuhan yang serupa. Strategi ini memainkan peran vital dalam dunia bisnis karena dapat menyediakan informasi penting mengenai loyalitas pelanggan serta membantu perusahaan dalam menyusun strategi pemasaran yang lebih efisien dan tertarget.

Sejumlah penelitian sebelumnya menunjukkan bahwa teknik data mining telah banyak dimanfaatkan dalam proses segmentasi pelanggan untuk mendukung pengambilan keputusan yang lebih tepat. Salah satu teknik yang sering digunakan dalam proses ini adalah clustering, yaitu metode pengelompokan data berdasarkan kemiripan tertentu. Pada studi ini, dilakukan analisis perbandingan terhadap dua algoritma

e-ISSN: 2775 – 8672
p-ISSN: 2775 – 9482

clustering yang populer, yaitu K-Means dan K-Medoids, dalam konteks segmentasi pelanggan pada platform e-commerce.

Algoritma K-Means merupakan metode clustering berbasis partisi yang mengelompokkan data ke dalam k klaster berdasarkan kemiripan fitur. Setiap data akan bergabung ke klaster yang memiliki pusat (centroid) paling dekat dengannya. Dengan demikian, data dalam satu klaster memiliki kesamaan yang tinggi, namun berbeda secara signifikan dari data dalam klaster lainnya.

Sebaliknya, algoritma K-Medoids juga termasuk metode clustering berbasis partisi, namun menggunakan titik aktual dari data (medoid) sebagai pusat klaster, bukan nilai rata-rata seperti pada K-Means. Medoid dianggap sebagai titik yang paling representatif dalam klaster dan cenderung lebih tahan terhadap outlier dibandingkan dengan centroid milik K-Means.

Untuk mengkaji kelebihan dan kekurangan dari kedua algoritma tersebut, penelitian ini menggunakan pendekatan Systematic Literature Review (SLR) guna meninjau dan menganalisis sejumlah studi terdahulu yang membahas penerapan K-Means dan K-Medoids dalam proses segmentasi pelanggan e-commerce. Pendekatan ini memungkinkan peneliti menyusun temuan secara objektif, sistematis, dan menyeluruh berdasarkan berbagai sumber primer dalam lima tahun terakhir. Diharapkan, melalui sintesis literatur ini, penelitian mampu memberikan kontribusi teoritis maupun praktis dalam memilih algoritma segmentasi pelanggan yang paling sesuai dengan karakteristik data serta kebutuhan bisnis saat ini.

METODE

Penelitian ini mengadopsi pendekatan Systematic Literature Review (SLR) sebagai metode utama untuk menelaah serta membandingkan penggunaan algoritma K-Means dan K-Medoids dalam proses segmentasi pelanggan pada sektor e-commerce. Metode SLR dipilih karena mampu menyajikan rangkuman informasi yang menyeluruh, sistematis, dan tidak bias dari berbagai penelitian sebelumnya yang relevan dengan topik ini. Proses pelaksanaan SLR dilakukan melalui beberapa tahapan sebagai berikut: 1) Perumusan Pertanyaan Penelitian. Fokus utama penelitian ini adalah untuk mengevaluasi perbandingan kinerja antara algoritma K-Means dan K-Medoids dalam konteks segmentasi pelanggan pada platform e-commerce; 2) Tahap Pengumpulan Literatur. Setelah topik ditetapkan, penelusuran terhadap artikel-artikel ilmiah yang relevan dilakukan dengan menggunakan platform Google Scholar (Google Cendekia), dengan cakupan waktu publikasi lima tahun terakhir. Kata kunci yang digunakan berkaitan dengan implementasi K-Means dan K-Medoids dalam segmentasi data pelanggan. Dari proses ini, berhasil dikumpulkan 12 artikel yang dinilai layak dan sesuai untuk dianalisis lebih lanjut dalam kajian sistematis; 3) Pengolahan dan Analisis Data. Literatur yang telah diperoleh kemudian dikaji secara mendalam guna memperoleh informasi penting yang berhubungan dengan metode, hasil temuan, serta kelebihan dan keterbatasan dari kedua algoritma yang diteliti. Informasi yang terkumpul disusun dalam format matriks analisis untuk mempermudah proses sintesis dan perbandingan antar studi; 4) Penyusunan Kesimpulan. Langkah terakhir adalah menyusun kesimpulan berdasarkan hasil analisis yang telah dilakukan. Kesimpulan ditarik dengan cara mengidentifikasi hubungan dan pola informasi dari studi-studi yang telah direview, sehingga menghasilkan pemahaman baru terkait efektivitas penerapan algoritma K-Means dan K-Medoids dalam segmentasi pelanggan e-commerce.

HASIL DAN PEMBHASAN

Ringkasan Studi yang Direview

Dari 12 jurnal yang dikaji, seluruhnya menggunakan algoritma K-Means dan/atau K-Medoids dalam konteks segmentasi pelanggan *e-commerce*. Beberapa studi juga membandingkan performa kedua algoritma dengan mempertimbangkan metrik evaluasi seperti SSE *(Sum of Squared Error), Silhouette Score*, waktu komputasi, dan stabilitas hasil klaster. Tabel berikut merangkum temuan utama dari setiap jurnal:

Tabel 1. Tabel *Review* Jurnal

Referensi	Dataset	Algoritma	Metrik Evaluasi	Hasil Utama
Perdana, Florentin, & Santoso, 2022	Data Transaksi Alfagift	K-Means	SSE, *Elbow*	K=3 (*Elbow*, SSE=1119.29); hasil segmentasi digunakan untuk strategi pemasaran.
Luchia Awalina & Rahayu, 2023	Data Transaksi Onine Retail	K-Means	*Elbow*	K=4 (*Elbow*); klaster terbentuk berdasarkan *quantity*, unit *price*, dan *customer* ID; kuantitas dan harga satuan memengaruhi perilaku pelanggan.
Dwi Sulistyawati & Sadikin, 2021	Data Transaksi Penjualan	K-Medoids	*Silhouette Index dan Davies*	Jumlah klaster optimal adalah 3 dengan Silhouette Index 0,375 dan *Davies Bouldin Index* 1,030, menghasilkan segmen pelanggan: lost customer, core customer, dan new customer.

e-ISSN: 2775 – 8672
p-ISSN: 2775 – 9482

Ardana, Adlian Aldita Alif, & Faisal, 2024	Data transaksi dari toko-toko online/retail di UK	K-Means	SSE	*Cluster* 0 memiliki total transaksi tertinggi, sedangkan *Cluster* 1 adalah pelanggan dengan frekuensi transaksi paling sering.
Ardi, Nastiti, & Sumarlinda, 2023	Data Transaksi Penjualan	K-Means	*Silhouette Scores*	Hasil penelitian menunjukkan 2 *cluster* pelanggan: *Cluster* 1 dengan 343 anggota dan *Cluster* 2 dengan 8 anggota. Metode *Silhouette Score* menegaskan bahwa 2 *cluster* adalah jumlah paling optimal.
Huda, Fatimah, & Rachmawati, 2022	Data Nasabah dan Data Transaksi Nasabah	K-Means	SSE	Hasil k-means *clustering* dari 1000 data: *Cluster* 1 (286 nasabah/28,6%), *Cluster* 2 (130/13%), *Cluster* 3 (542/54,2%), dan *Cluster* 4 (42/4,2%).
Pangestu, Hermanto, & Irmayanti, 2023	Data Transaksi Pelanggan Perusahaan	K-Means	*Silhouette Coefficient*	Hasil penelitian: terbentuk 3 *cluster* yaitu *Low Loyalty* (137 pelanggan), *Medium Loyalty* (1.636), dan *Highest Loyalty* (2.395).
Mirantika, Syamfithriani, & Trisudarno, 2023	Data Transaksi Penjualan Produk	K-Medoids	*Davies Bouldin Index*	Hasil segmentasi: 3 segmen pelanggan – Loyal Customer (21, recency tinggi, frequency & monetary tertinggi), Typical Customer (31, recency rendah, frequency & monetary sedang), Newcomer (61, recency tertinggi, frequency & monetary terendah).
Windjaya & Siregar, 2024	Database Salah Satu Toko Online Marketplace di Indonesia	K-Medoids	*Davies Bouldin Index*	K-Medoids RFMTM menghasilkan 4 *cluster* dengan karakteristik berbeda; *cluster* terbaik pelanggan loyal, terburuk pelanggan tidak aktif.
Rohman & Wibowo, 2024	Data penelitian menggunakan *dataset* publik dari kaggle.com	K-Means, K-Medoids	*Silhouette Coefficient*	Nilai Silhouette Coefficient k-means 0,553 dan k-medoids 0,485; algoritma terbaik adalah k-means karena memiliki nilai tertinggi.
Murpratiwi, Hendrawan, & Aranta, 2021	Data Transaksi Toko Retail	K-Means, K-Medoids	*Davies Bouldin Index*	K-Medoids memiliki validitas terbaik (DBI 0,540778) dibanding K-Means & K-Means; jumlah *cluster* optimal adalah 5.
Madani, Rahmah, Nurunnisa, & Elia, 2022	Data Transaksi Pelanggan	K-Means	*Davies Bouldin Index*	Diperoleh 2 *cluster* terbaik (DBI 0,492) dengan segmen 452 core customer dan 351 lost customer.
Widiyanto & Witanti, 2021	Data Pelanggan	K-Means	*Davies Bouldin Index*	Hasil: Terbentuk 4 segmen pelanggan; dari 29 pelanggan, akurasi sistem 100% sesuai pengetahuan user; membantu PT Coversuper memprioritaskan pelanggan potensial.

Berdasarkan tinjauan literatur yang ditemukan, algoritma K-Means tercatat lebih banyak digunakan dibandingkan dengan K-Medoids dalam penerapan segmentasi pelanggan, karena dianggap lebih efisien dalam mengolah data berukuran besar dan mampu menghasilkan pembagian *cluster* yang optimal pada berbagai studi kasus.

KESIMPULAN

Berdasarkan studi literatur yang membandingkan algoritma K-Means dan K-Medoids dalam proses segmentasi pelanggan, dapat disimpulkan bahwa kedua metode memiliki kelebihan masing-masing dalam menghasilkan pembagian *cluster* yang sesuai dengan karakteristik data. Sebagian besar penelitian menunjukkan bahwa K-Means lebih sering digunakan karena mampu menangani *dataset* berukuran besar dengan waktu komputasi yang relatif cepat serta menghasilkan *cluster* yang jelas, seperti pada studi dengan nilai Silhouette Coefficient 0,553 dan berbagai hasil segmentasi yang memunculkan pola pelanggan seperti *lost customer*, *core customer*, *new customer*, hingga segmentasi berdasarkan tingkat loyalitas. Sementara itu, K-Medoids menunjukkan performa lebih baik pada data yang mengandung *Outlier* serta memberikan nilai validitas yang stabil, seperti pada penelitian dengan Davies Bouldin Index (DBI) 0,540778 yang mengindikasikan kualitas *cluster* lebih baik dibandingkan X-Means dan K-Means pada kasus tertentu. Beberapa studi juga menegaskan bahwa K-Medoids lebih sesuai ketika *dataset* berukuran kecil atau saat diperlukan ketahanan terhadap noise. Secara keseluruhan, literatur menunjukkan bahwa K-Means unggul dalam efisiensi dan interpretasi hasil untuk segmentasi pelanggan skala besar, sedangkan K-Medoids lebih andal dalam menjaga ketepatan *cluster* pada data yang kompleks atau memiliki variasi tinggi. Kombinasi hasil ini memberikan dasar penting bagi peneliti dan praktisi untuk memilih algoritma yang sesuai dengan karakteristik data dan tujuan analisis dalam strategi segmentasi pelanggan.

e-ISSN: 2775 – 8672
p-ISSN: 2775 – 9482

DAFTAR REFERENSI

Ardi, R. B., Nastiti, F. E., & Sumarlinda, S. (2023). Algoritma K-Means *clustering* untuk segmentasi pelanggan (studi kasus fashion viral Solo). *Infotech Journal*, *1*, 124–132.

Ardana, C. H., Alif, A. A. A., & Faisal, M. (2024). Segmentasi pelanggan penjualan online menggunakan metode K-Means *clustering*. *JISKA (Jurnal Informatika Sunan Kalijaga)*, *1*(1), 1–9.

Dwi Sulistyawati, A. A., & Sadikin, M. (2021). Penerapan algoritma K-Medoids untuk menentukan segmentasi pelanggan. *SISTEMASI: Jurnal Sistem Informasi*, *10*(3), 516–526.

Evanita, F. M., Cholissodin, I., & Adinugroho, S. (2021). Pengelompokan toko *e-commerce* Shopee berdasarkan reputasi toko menggunakan metode *clustering* K-Medoids. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *5*(4), 1230–1233.

Huda, N. A., Fatimah, F., & Rachmawati, F. (2022). Analisis segmentasi pelanggan bank berdasarkan pengambilan kredit dengan menggunakan metode K-Means *clustering*. *Jurnal Ilmiah Teknologi Informasi Terapan*, *6*(2), 393–400.

Kamila, C., A. S., M., Namang, G. R., & Fadila Syah, R. R. (2021). *Systematic literature review*: Penggunaan algoritma K-Means untuk *clustering* di Indonesia dalam bidang pendidikan. *Jurnal Informatika dan Teknologi*, *7*(1), 101–110.

Luchia Awalina, E. F., & Rahayu, W. I. (2023). Optimalisasi strategi pemasaran dengan segmentasi pelanggan menggunakan penerapan K-Means *clustering* pada transaksi online retail. *Jurnal Teknologi dan Informasi (JAIT)*, *12*(1), 122–137.

Madani, A., Rahmah, A., Nurunnisa, F., & Elia, A. (2022). Segmentasi pelanggan pada BC HNI 2 Pekanbaru dengan menerapkan algoritma K-Medoids dan model RFM (*Recency, Frequency, Monetary*). *SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat*, *6*, 179–186.

Mirantika, N., Syamfithriani, T. S., & Trisudarno, R. (2023). Implementasi algoritma K-Medoids *clustering* untuk menentukan segmentasi pelanggan. *Jurnal Nuansa Informatika*, *9*(2), 196–204.

Murpratiwi, S. I., Hendrawan, I. A., & Aranta, A. (2021). Analisis pemilihan *cluster* optimal dalam segmentasi pelanggan toko retail. *Jurnal Pendidikan Teknologi dan Kejuruan*, *14*(2), 152–163.

Olivia, L. F., Juliantho, D. A., & Hendrik, B. (2023). Komparasi algoritma K-Means dan K-Medoids dalam *clustering* penyebaran kasus COVID-19. *Journal of Information System and Education Development*, *3*(2), 30–31.

Pangestu, P. I., Hermanto, T. I., & Irmayanti, D. (2023). Analisis segmentasi pelanggan berbasis *Recency, Frequency, Monetary* (RFM) menggunakan algoritma K-Means. *JAIT (Jurnal Mahasiswa Teknik Informatika)*, *8*(3), 1486–1492.

Perdana, S. A., Florentin, S. F., & Santoso, A. (2022). Analisis segmentasi pelanggan menggunakan K-Means *clustering*: Studi kasus aplikasi Alfagift. *Sebatik*, *26*(3), 446–448.

Rohman, N., & Wibowo, A. (2024). Perbandingan metode K-Medoids dan metode K-Means dalam analisis segmentasi pelanggan mall. *SINTECH: Science and Information Technology Journal*, *5*(1), 49–58.

Widiyanto, A., & Witanti, A. (2021). Segmentasi pelanggan berdasarkan analisis RFM menggunakan algoritma K-Means sebagai dasar strategi pemasaran (studi kasus PT Coversuper Indonesia Global). *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, *9*(2), 204–215.

Windjaya, P. A., & Siregar, B. (2024). Analisis segmentasi pelanggan toko online marketplace berdasarkan *Recency, Frequency, Monetary, Time, Satisfaction* (*RFMTMS*) menggunakan algoritma K-Medoids *clustering*. *Mutiara: Multidisciplinary Scientific Journal*, *10*(1), 888–906.