BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

Segmentasi Pelanggan WiFi RT/RW Net Menggunakan Metode K-Means Clustering untuk Identifikasi Potensi Churn sebagai Strategi Retensi Pelanggan

Customer Segmentation for RT/RW Net WiFi Services Using the K-Means Clustering Method to Identify Potential Churn as a Customer Retention Strategy

Mochamad David Hariono*, Haris Yuana, Wahyu Dwi Puspitasari

Fakultas Teknik dan Informatika, Program Studi Teknik Informatika, Universitas Islam Balitar, Blitar, Indonesia
Email: [1,*]mochamaddavidhariono@gmail.com, [2]harisyuana2010@gmail.com, [3]pusphitasari23@gmail.com
Email Penulis Korespondensi: mochamaddavidhariono@gmail.com
Received: **04/07/2026**, Revised: **30/10/2026**, Accepted: **31/10/2026**, Published**: 31/10/2026**

Abstrak−Perkembangan teknologi informasi menjadikan internet sebagai kebutuhan utama masyarakat dalam menunjang berbagai aktivitas sehari-hari. Meningkatnya persaingan antarpenyedia layanan internet menjadikan retensi pelanggan sebagai salah satu tantangan penting dalam menjaga keberlanjutan bisnis, khususnya dalam menghadapi fenomena *churn*, yaitu kondisi ketika pelanggan menghentikan penggunaan layanan dan beralih ke penyedia lain. Penelitian ini bertujuan melakukan segmentasi pelanggan untuk mengidentifikasi potensi *churn* menggunakan metode *K-Means Clustering* pada 984 pelanggan aktif CV. Bagus Agung Mandiri Sejahtera. Variabel yang digunakan meliputi lama berlangganan, frekuensi kedisiplinan pembayaran, kepuasan terhadap layanan teknisi dan *customer service,* serta niat berpindah layanan. Pengolahan data dilakukan melalui tahapan *Knowledge Discovery in Database* (KDD). Hasil penelitian menunjukkan terbentuknya tiga klaster, yaitu klaster 1 sebanyak 776 pelanggan (78,86%), klaster 2 sebanyak 161 pelanggan (16,36%), dan klaster 3 sebanyak 47 pelanggan (4,78%). Berdasarkan hasil *profiling*, klaster 2 terindikasi sebagai kelompok dengan indikasi potensi *churn* tertinggi, yang ditandai dengan nilai lama berlangganan sebesar 0,1801, frekuensi kedisiplinan pembayaran sebesar 0,1843, kepuasan teknisi sebesar 0,207, kepuasan *customer service* sebesar 0,1946, serta niat berpindah layanan sebesar 0. Karakteristik tersebut menunjukkan bahwa pelanggan pada klaster 2 cenderung memiliki masa berlangganan lebih singkat, kedisiplinan pembayaran rendah, kepuasan layanan rendah, dan kecenderungan berpindah layanan tinggi. Evaluasi menggunakan *Silhouette Coefficient* menghasilkan nilai sebesar 0,5773, yang menunjukkan bahwa struktur klaster memiliki kualitas pemisahan yang baik berdasarkan karakteristik data pelanggan. Hasil penelitian ini dapat menjadi dasar penyusunan strategi retensi yang lebih terarah, khususnya melalui peningkatan kualitas pelayanan teknisi dan *customer service*, pemberian pengingat pembayaran secara berkala, serta pendekatan personal dan penawaran program loyalitas kepada pelanggan pada klaster 2.

Kata Kunci: Segmentasi Pelanggan; K-Means Clustering; Churn Pelanggan; Knowledge Discovery in Database (KDD); Silhouette Coefficient

Abstract−Advances in information technology have made the internet an essential need for society in supporting various daily activities. Increasing competition among internet service providers has made customer retention one of the key challenges in ensuring business sustainability, particularly in addressing the phenomenon of churn that is, when customers stop using a service and switch to another provider. This study aims to segment customers to identify potential churn using the K-Means Clustering method on 984 active customers of CV. Bagus Agung Mandiri Sejahtera. The variables used include subscription duration, payment discipline frequency, satisfaction with technician and customer service, and intention to switch services. Data processing was conducted through the Knowledge Discovery in Databases (KDD) methodology. The results showed the formation of three clusters: Cluster 1 comprising 776 customers (78.86%), Cluster 2 comprising 161 customers (16.36%), and Cluster 3 comprising 47 customers (4.78%). Based on the profiling results, Cluster 2 is identified as the group with the highest potential for churn, characterized by a subscription duration value of 0.1801, a payment discipline frequency of 0.1843, technician satisfaction of 0.207, customer service satisfaction of 0.1946, and an intention to switch providers of 0. These characteristics indicate that customers in Cluster 2 tend to have shorter subscription periods, low payment discipline, low service satisfaction, and a high tendency to switch services. An evaluation using the Silhouette Coefficient yielded a value of 0.5773, indicating that the cluster structure provides good separation based on customer data characteristics. The results of this study can serve as a basis for developing more targeted retention strategies, particularly through improving the quality of technician and customer service, sending periodic payment reminders, and offering personalized approaches and loyalty programs to customers in Cluster 2.

Keywords: Customer Segmentation; K-Means Clustering; Customer Churn; Knowledge Discovery in Databases (KDD); Silhouette Coefficient

1. PENDAHULUAN

Perkembangan teknologi informasi menjadikan internet sebagai salah satu kebutuhan utama masyarakat dalam menunjang berbagai aktivitas, seperti pendidikan, pekerjaan, komunikasi, dan kegiatan ekonomi. Berdasarkan laporan Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) tahun 2025, tingkat penetrasi internet di Indonesia mencapai 80,66% [1]. Meskipun demikian, pemerataan akses internet masih menjadi tantangan, terutama pada wilayah yang belum sepenuhnya terjangkau oleh infrastruktur jaringan penyedia layanan internet berskala besar [2]. Kondisi tersebut mendorong berkembangnya penyedia layanan internet lokal, salah satunya melalui model bisnis RT/RW Net. RT/RW Net merupakan layanan distribusi akses internet yang umumnya dilakukan dengan membeli *bandwidth* dari penyedia

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

layanan internet utama atau *Internet Service Provider* (ISP), kemudian mendistribusikannya kembali kepada pelanggan di wilayah tertentu [3]. Keberadaan penyedia layanan internet lokal memberikan akses internet bagi masyarakat, tetapi juga menghadapi tantangan dalam mempertahankan pelanggan di tengah persaisangan layanan yang semakin kompetitif.

Salah satu tantangan yang dihadapi oleh penyedia layanan internet adalah fenomena *customer churn*, yaitu kondisi ketika pelanggan menghentikan penggunaan layanan dan beralih ke penyedia lain. Fenomena tersebut dapat berdampak terhadap keberlangsungan bisnis karena perusahaan tidak hanya kehilangan pelanggan, tetapi juga berpotensi mengalami penurunan pendapatan serta peningkatan biaya untuk memperoleh pelanggan baru [4]. Permasalahan tersebut juga dihadapi oleh CV. Bagus Agung Mandiri Sejahtera sebagai salah satu penyedia layanan internet berbasis RT/RW Net. Berdasarkan data perusahaan, selama tahun 2025 tercatat 126 pelanggan berhenti berlangganan. Kondisi ini menunjukkan perlunya upaya yang lebih terarah dalam memahami karakteristik pelanggan agar perusahaan dapat mengidentifikasi kelompok pelanggan yang memiliki indikasi potensi *churn*. Selama ini, data pelanggan yang tersedia belum dimanfaatkan secara optimal untuk mengelompokkan pelanggan berdasarkan karakteristik lama berlangganan, perilaku pembayaran, dan persepsi terhadap kualitas layanan. Oleh karena itu, diperlakukan pendekatan analisis berbasis data yang dapat membantu perusahaan memperoleh informasi mengenai karakteristik pelanggan sebagai bahan pertimbangan dalam menentukan prioritas strategi retensi.

Salah satu pendekatan yang dapat digunakan untuk memperoleh informasi dari data pelanggan adalah *Knowledge Discovery in Database* (KDD). KDD merupakan proses sistematis untuk menemukan pengetahuan yang bermanfaat dari sekumpulan data melalui beberapa tahapan, yaitu seleksi data (*selection*), prapemrosesan (*preprocessing*), transformasi (*transformation*), penerapan teknik data mining, serta interpretasi dan evaluasi hasil [5]. Penggunaan kerangka KDD memungkinkan proses pengolahan data dilakukan secara terstruktur, mulai dari penentuan variabel yang relevan hingga penarikan informasi berdasarkan hasil analisis. Pada tahap data mining, salah satu metode yang dapat diterapkan untuk mengelompokkan pelanggan berdasarkan kemiripan karakteristik adalah *K-Means Clustering*. *K-Means* merupakan algoritma klasterisasi partisional yang membagi sejumlah objek ke dalam kelompok berdasarkan kedekatan jarak terhadap pusat klaster (*centroid*) [6]. Prosesnya dilakukan dengan menentukan jumlah klaster, menetapkan *centroid* awal, menghitung jarak setiap objek terhadap *centroid,* menempatkan objek ke klaster dengan jarak terdekat, kemudian memperbarui *centroid* berdasarkan rata-rata anggota klaster secara iterative hingga memenuhi kondisi penghentian (*convergence*) [7]. Dalam penelitian ini, *K-Means* diterapkan untuk mengelompokkan pelanggan berdasarkan lama berlangganan, frekuensi kedisiplinan pembayaran, kepuasan terhadap layanan teknisi, kepuasan terhadap *customer service*, dan niat berpindah layanan.

Pemilihan *K-Means* dalam penelitian ini juga mempertimbangkan karakteristik dan kebutuhan analisis segmentasi pelanggan. Terdapat beberapa algoritma klasterisasi lain yang dapat digunakan, di antaranya *K-Medoids* dan *DBSCAN*. *K-Medoids* menggunakan objek aktual dalam dataset sebagai pusat klaster sehingga relative lebih tahan terhadap pengaruh pencilan (*outlier*) dibandingkan *K-Means* yang menggunakan nilai rata-rata sebagai *centroid* [8] . Sementara itu, *DBSCAN* menerapkan pendekatan berbasis kepadatan untuk membentuk kelompok berdasarkan kepadatan data dan mampu mengidentifikasi objek yang dikategorikan sebagai *noise* tanpa harus menentukan jumlah klaster sejak awal. Namun, *DBSCAN* memerlukan penentuan parameter *Epsilon* (ε) dan *MinPts* yang sesuai dengan karakteristik distribusi data [9]. Berdasarkan tujuan penelitian yang berorientasi pada pembentukan tiga kelompok pelanggan untuk keperluan *customer profiling*, *k-means* dipilih karena mekanisme pengelompokannya sesuai dengan kebutuhan segmentasi berdasarkan kedekatan karakteristik pelanggan. Jumlah tiga klaster ditetapkan berdasarkan kebutuhan analisis segmentasi dalam penelitian ini, bukan sebagai hasil pencarian jumlah klaster yang paling optimal. Meskipun demikian, *k-means* memiliki keterbatasan berupa sensitivitas terhadap pencilan karena penggunaan nilai rata-rata sebagai pusat klaster [10]. Dalam penelitian ini, pemeriksaan dan penanganan *outlier* belum dilakukan secara khusus sehingga pengaruh pencilan menjadi salah satu keterbatasan yang perlu diperhatikan dalam menginterpretasikan hasil pengelompokan dan dapat dikaji lebih lanjut pada penelitian berikutnya.

Untuk mengetahui kualitas hasil pengelompokan, diperlukan proses evaluasi terhadap struktur klaster yang terbentuk. Salah satu metode evaluasi yang dapat digunakan adalah *Silhouette Coefficient*, yaitu metode yang mengukur kualitas klaster berdasarkan tingkat kekompakan objek dalam klaster yang sama (*cohesion*) dan tingkat keterpisahan terhadap klaster lainnya (*separation*) [11]. Nilai *silhouette* berada pada rentang -1 hingga 1. Nilai yang mendekati 1 menunjukkan bahwa objek memiliki posisi yang semakin sesuai dengan klasternya, nilai mendekati 0 menunjukkan adanya kemungkinan tumpang tindih antar klaster, sedangkan nilai negative menunjukkan bahwa objek berpotensi lebih sesuai ditempatkan pada klaster lain [7], [4]. Penggunaan *silhouette coefficient* dalam penelitian ini bertujuan untuk mengevaluasi kualitas struktur tiga klaster yang telah ditetapkan, khususnya dalam menilai tingkat kekompakan dan keterpisahan kelompok pelanggan. Dengan demikian, evaluasi tersebut digunakan untuk memberikan gambaran mengenai kualitas hasil segmentasi, bukan untuk membuktikan bahwa tiga klaster merupakan jumlah klaster yang paling optimal.

Beberapa penelitian terdahulu mengenai segmentasi dan fenomena *churn* pada industry telekomunikasi telah menerapkan teknik klasterisasi untuk mendukung pengambilan keputusan bisnis. Penelitian yang dilakukan oleh Julian dkk. (2025) pada sektor telekomunikasi berhasil mengelompokkan pelanggan ke dalam tiga klaster, yaitu pelanggan dengan kebutuhan layanan dasar, pelanggan yang sensitif terhadap harga serta pelanggan dengan intensitas penggunaan internet yang tinggi namun jarang memanfaatkan fitur tambahan [12]. Hasil segmentasi tersebut masih digunakan sebagai dasar dalam penyusunan strategi pemasaran berupa penawaran paket hemat, promosi *bundling*, dan edukasi mengenai manfaat fitur tambahan layanan. Penelitian R.N Rahmadiana & D.Lestarini (2025) menunjukkan bahwa loyalitas

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2706
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

pelanggan dipengaruhi oleh kualitas layanan dan tingkat kepuasan [13]. Meskipun dilakukan pada sektor perbankan, penelitian tersebut menunjukkan pentingnya mempertimbangkan aspek kualitas layanan dalam memahami perilaku pelanggan. Penelitian oleh Gunawan (2026) pada penyedia layanan internet menggunakan pendekatan *Stuctural Equation Modeling-Partial Least Square* (SEM-PLS) menunjukkan bahwa *customer loyalty* memiliki peran penting dalam memengaruhi keputusan pelanggan untuk tetap menggunakan layanan [14].

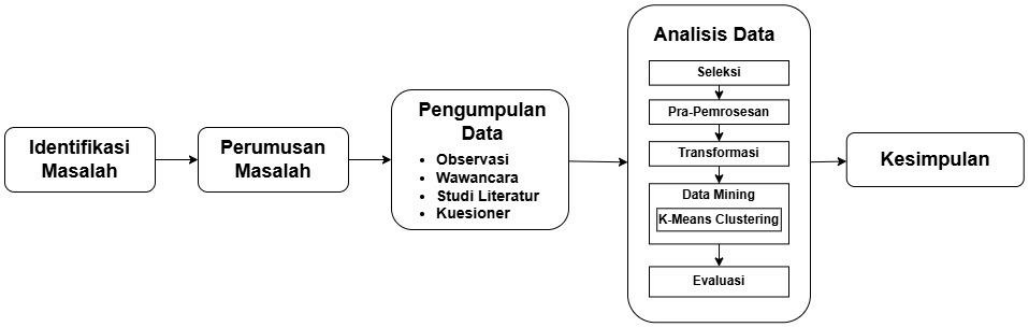
Penelitian yang dilakukan oleh Didik Priyanto dkk. (2020) mengidentifikasi bahwa faktor perilaku penggunaan layanan, seperti status *suspend*, metode pembayaran, serta peningkatan dan penurunan paket layanan, berkaitan dengan fenomena *churn* pelanggan. Sementara itu, penelitian yang dilakukan oleh Harani dkk (2020) menunjukkan bahwa analisis karakteristik pelanggan dapat dimanfaatkan untuk membangun *customer profiling* dan segemntasi pelanggan yang lebih akurat [15]. Kajian mengenai *churn* pada penyedia layanan internet juga dilakukan oleh Salma dan Aprianingsih (2021) yang menganalisis faktor-faktor penentu perpindahan pelanggan pada salah satu perusahaan ISP di Indonesia. Penelitian tersebut menyoroti persaingan antarpenyedia layanan internet dan penawaran kompetitor sebagai salah satu tantangan dalam mempertahankan pelanggan [16].

Berdasarkan penelitian terdahulu, kajian mengenai *churn* pada industri telekomunikasi dan penyedia layanan internet telah membahas berbagai faktor yang berkaitan dengan keputusan pelanggan untuk berhenti berlangganan atau berpindah penyedia layanan. Namun, masih diperlukan pendekatan segmentasi yang mengintegrasikan karakteristik perilaku pelanggan dan persepsi terhadap kualitas pelayanan ke dalam satu proses pengelompokan untuk mengidentifikasi kelompok pelanggan berdasarkan indikasi potensi *churn*, khususnya pada penyedia layanan internet skala lokal. Penelitian sebelumnya cenderung berfokus pada identifikasi faktor yang memengaruhi *churn*, pengujian hubungan antarvariabel, maupun analisis karakteristik penggunaan layanan. Berbeda dengan pendekatan tersebut, penelitian ini menerapkan *K-Means Clustering* untuk mengintegrasikan lima variabel, yaitu lama berlangganan, kedisiplinan pembayaran, kepuasan terhadap layanan teknisi, kepuasan terhadap *customer service*, dan niat berpindah layanan, ke dalam proses segmentasi pelanggan. Hasil pengelompokan selanjutnya dievaluasi menggunakan *Silhouette Coefficient* untuk mengetahui kualitas struktur klaster yang terbentuk. Melalui pendekatan tersebut, penelitian ini bertujuan menghasilkan segmentasi pelanggan berdasarkan kesamaan karakteristik serta mengidentifikasi kelompok yang memiliki indikasi potensi *churn* tertinggi. Hasil penelitian diharapkan dapat menjadi dasar bagi perusahaan dalam menentukan prioritas dan menyusun strategi retensi pelanggan secara lebih terarah.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menerapkan metode *Knowledge Discovery in Database* (KDD) karena menyediakan alur analisis data yang sistematis dan terstruktur mulai dari pengumpulan data hingga interpretasi hasil. Sebelum memasuki tahapan KDD, penelitian diawali dengan perumusan masalah dan pengumpulan data sebagai dasar dalam menentukan ruang lingkup penelitian serta memperoleh data yang akan dianalisis. Selanjutnya, proses KDD meliputi tahapan seleksi data, pra-pemrosesan, transformasi data, penerapan metode *k-means clustering*, serta evaluasi dan interpretasi hasil pengelompokan. Keseluruhan tahapan penelitian ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

a. Identifikasi Masalah
Tahap identifikasi masalah dilakukan untuk memperoleh pemahaman terhadap kondisi yang dihadapi oleh objek penelitian. Identifikasi dilakukan melalui observasi terhadap aktivitas operasional perusahaan, wawancara dengan pihak pengelola, serta analisis terhadap data pelanggan yang tersedia. Hasil identifikasi menunjukkan bahwa CV. Bagus Agung Mandiri Sejahtera sebagai penyedia layanan WiFi RT/RW Net menghadapi permasalahan berupa meningkatnya pelanggan yang menghentikan penggunaan layanan (*churn*) dan beralih ke penyedia layanan internet lainnya. Di sisi lain, data pelanggan yang dimiliki perusahaan belum dimanfaatkan secara optimal untuk mengidentifikasi karakteristik pelanggan maupun mendukung penyusunan strategi retensi yang tepat sasaran. Berdasarkan kondisi tersebut, penelitian ini menerapkan metode *K-Means Clustering* untuk melakukan segmentasi pelanggan berdasarkan karakteristik yang dimiliki sehingga dapat mengidentifikasi kelompok pelanggan yang berpotensi mengalami *churn* sebagai dasar dalam penyusunan strategi retensi pelanggan.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

b. Perumusan Masalah
Tahap perumusan masalah dilakukan berdasarkan hasil identifikasi terhadap kondisi perusahaan. Permasalahan penelitian difokuskan pada penerapan metode *K-Means Clustering* untuk melakukan segmentasi pelanggan berdasarkan karakteristik yang dimiliki sehingga dapat mengidentifikasi kelompok pelanggan yang berpotensi mengalami *churn*.

c. Pengumpulan Data
Tahap pengumpulan data dilakukan untuk memperoleh informasi yang diperlukan dalam proses segmentasi pelanggan. Data penelitian dikumpulkan melalui observasi, wawancara, studi literatur, dan penyebaran kuesioner. Observasi dan wawancara dilakukan kepada pihak CV. Bagus Agung Mandiri Sejahtera untuk memperoleh informasi mengenai kondisi operasional perusahaan, karakteristik layanan, serta faktor-faktor yang berkaitan dengan terjadinya *churn* pelanggan. Selanjutnya, studi literatur dilakukan dengan menelaah buku, jurnal, dan referensi ilmiah yang relevan sebagai landasan teoritis dalam penelitian. Selain itu, penyebaran kuesioner kepada pelanggan digunakan untuk memperoleh data mengenai persepsi pelanggan terhadap kualitas layanan, yang meliputi kepuasan terhadap layanan teknisi, kepuasan layanan *customer service*, serta niat berpindah layanan. Seluruh data yang diperoleh pada tahap ini menjadi dasar dalam proses segmentasi pelanggan menggunakan metode *K-Means Clustering*.

d. Analisis data
Analisis data pada penelitian ini mengikuti tahapan *Knowledge Discovery in Database* (KDD) [5] yaitu: (1) Seleksi data dilakukan untuk menentukan variabel yang digunakan dalam proses analisis sesuai dengan tujuan penelitian; (2) Pra-pemrosesan yaitu menyiapkan data dengan melakukan pemeriksaan terhadap ketidakkonsistenan, perbaikan kesalahan, serta penambahan informasi untuk meningkatkan kualitas data; (3) Transformasi dilakukan dengan pemberian skala pada setiap variabel untuk mempermudah proses analisis; (4) Tahap *data mining* menerapkan metode *k-means clustering* untuk mengelompokkan pelanggan berdasarkan karakteristik yang dimiliki sehingga diperoleh pola segmentasi yang dapat mengidentifikasi potensi churn; (5) Evaluasi dilakukan untuk menilai kualitas hasil pengelompokan menggunakan *silhouette coefficient*, sehingga dapat dikatehui tingkat kedekatan dan pemisahan antar klaster yang terbentuk.

e. Kesimpulan
Kesimpulan dilakukan dengan merangkum seluruh temuan yang diperoleh selama proses penelitian serta mengaitkannya dengan tujuan penelitian yang telah ditetapkan. Kesimpulan yang dihasilkan menggambarkan hasil penerapan metode *k-means clustering* dapat mengidentifikasi potensi *churn* melalui segementasi pelanggan.

2.2 Implementasi Algoritma K-Means Clustering

a. Menentukan Jumlah Klaster
Penelitian ini menggunakan tiga klaster yang ditetapkan berdasarkan kebutuhan analisis untuk membentuk segmentasi pelanggan ke dalam kelompok dengan karakteristik yang berbeda yaitu pelanggan loyal, rentan, dan beresiko *churn*. Penetapan jumlah klaster tersebut tidak didasarkan pada pengujian optimasi jumlah klaster menggunakan metode *Elbow*. Oleh karena itu, tiga klaster yang digunakan dalam penelitian ini merupakan jumlah kelompok yang ditetapkan untuk kebutuhan segmentasi, bukan hasil yang dapat dinyatakan sebagai jumlah klaster paling optimal secara statistik.

b. Menentukan Titik Pusat Awal *Cluster*
Centroid dipilih sebagai representasi awal dari masing-masing klaster dan digunakan sebagai acuan dalam proses pengelompokan data. Pada penelitian ini, *centroid* awal dipilih dari data pelanggan yang memiliki jarak relatif berjauhan sehingga setiap *centroid* mampu merepresentasikan karakteristik awal klaster yang berbeda dan menghasilkan pemisahan klaster yang lebih baik.

c. Menghitung Jarak *Centroid*
Setiap data pelanggan dihitung jaraknya terhadap seluruh *centroid* menggunakan metode *Euclidean Distance*. Perhitungan ini bertujuan untuk mengetahui tingkat kedekatan setiap data dengan masing-masing *centroid* sehingga dapat ditentukan klaster yang paling sesuai [17]. Perhitungan tersebut ditunjukkan pada Persamaan (1).

$$d(xi, \mu j) = \sqrt{\sum(xi - \mu j)^2} \quad (1)$$

Dengan $d(xi,\mu j)$ merupakan nilai jarak *Euclidean* antara data ke-*i* dan centroid pada klaster ke-*j, xi* menyatakan data ke-*i* yang akan dikelompokkan, dan μj sebagai *centroid* pada klaster ke-*j*.

d. Mengelompokkan Objek Berdasarkan Nilai Minimum
Hasil dari perhitungan jarak terhadap *centroid*, setiap data ditempatkan pada klaster yang memiliki nilai jarak terkecil (minimum). Proses ini diilakukan berdasarkan prinsip bahwa data akan menjadi anggota klaster yang *centroid* nya memiliki tingkat kedekatan paling tinggi.

e. Memperbarui Nilai *Centroid*
Setelah proses pengelompokkan selesai dilakukan, nilai *centroid* pada setiap klaster diperbarui dengan menghitung rata-rata seluruh data yang menjadi anggota klaster tersebut [17]. *Centroid* baru dihitung menggunakan Persamaan (2).

$$\mu j\ (t+1) = \frac{1}{Ns_j} \sum_j \epsilon\ s_j\ \ xj \quad (2)$$

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2708
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

Dengan $\mu j(t+1)$ merupakan *centroid* baru pada klaster ke-*j* pada iterasi ke-(t+1), Ns_j yaitu banyaknya data yang terdapat pada *cluster*, dan *xi* merupakan data pada klaster tersebut. Nilai *centroid* diperbarui dengan menghitung rata-rata seluruh data dalam setiap klaster. Proses iterasi kemudian diulangi hingga nilai *centroid* atau keanggotaan klaster tidak lagi mengalami perubahan (konvergen).

2.3 Evaluasi Silhouette Coefficient

a. Menghitung nilai *a*(*i*)
Nilai *a*(*i*) dihitung sebagai rata-rata jarak antara data ke-*i* dengan seluruh anggota pada klaster yang sama [11]. Persamaan yang digunakan ditunjukkan pada persamaan (3).

$$a(i) = \frac{1}{|A|-1}\sum_{j\in A, j\neq i} d(i,j) \quad (3)$$

Dengan *a*(*i*) adalah rata-rata jarak data ke-*i* terhadap seluruh data lain dalam klaster yang sama, $d(i,j)$ menyatakan jarak antara data ke-*i* dan data ke-j, sedangkan *A* merupakan klaster tempat data ke-*i* menjadi anggota.

b. Menghitung Nilai *b*(*i*)
Perhitungan dengan rata-rata objek terhadap semua data yang berada pada klaster lain menggunakan Persamaan 4.

$$d(i,C) = \frac{1}{|C|}\sum_{j\in C} d(i,j) \quad (4)$$

Dengan $d(i,C)$ merupakan rata-rata jarak antara data ke-*i* dan seluruh data pada klaster *C*, $d(i,j)$ menyatakan jarak antara data ke-*i* dan data ke-*j*, sedangkan |*C*| adalah jumlah data yang terdapat pada klaster *C*. Nilai *b*(*i*) kemudian diperoleh dengan memilih nilai $d(i,C)$ terkecil dari seluruh klaster selain klaster asal data ke-*I* [11], sebagaimana ditunjukkan pada Persamaan (5).

$$b(i) = min_{c\neq A}\, d(i,C) \quad (5)$$

Dimana *b*(*i*) adalah perbandingan dari klaster lain yang mencapai nilai *minimum*

c. Menghitung Nilai *s*(*i*)

$$s(i) = \frac{(bi-ai)}{\max(ai,bi)} \quad (6)$$

Dimana $s(i)$ adalah nilai akhir dari evaluasi, yang keseluruhan nilainya di rata-rata.

2.4 Normalisasi MinMax

Normalisasi data dalam penelitian ini dilakukan menggunakan metode *Min-Max* dengan mengubah rentang nilai setiap variabel ke dalam interval 0 hingga 1 [18]. Proses ini bertujuan menyamakan skala pengukuran antarvariabel sehingga setiap variabel dapat digunakan secara sebanding dalam perhitungan jarak pada algoritma *K-Means*. Meskipun demikian, normalisasi *Min-Max* tidak menghilangkan pengaruh pencilan (*outlier*) karena nilai ekstrem tetap dapat memengaruhi rentang minimum dan maksimum data. Dalam penelitian ini, pemeriksaan dan penanganan *outlier* belum dilakukan secara khusus sebelum proses klasterisasi. Oleh karena itu, kemungkinan pengaruh pencilan terhadap perhitungan *centroid* dan pembentukan klaster menjadi salah satu keterbatasan dalam penelitian ini. Proses perhitungan normalisasi *min-max* menggunakan persamaan 7.

$$X_n = \frac{x_0 - x_{min}}{x_{max} - x_{min}} \quad (7)$$

Dengan X_n merupakan nilai data hasil normalisasi, x_0 adalah nilai data aktual, x_{min} menyatakan nilai minimum dari seluruh data pada variabel yang sama, dan x_{max} merupakan nilai maksimum dari seluruh data pada variabel tersebut.

3. HASIL DAN PEMBAHASAN

3.1 Seleksi Data

Pada tahap seleksi data dilakukan pemilihan variabel yang relevan untuk digunakan dalam proses *clustering*. Variabel yang dipilih dari *database* perusahaan meliputi ID pelanggan, tanggal pemasangan, dan riwayat data pembayaran selama periode Januari hingga Desember 2025. ID pelanggan digunakan sebagai identitas setiap data pelanggan, sedangkan tanggal pemasangan digunakan untuk mengetahui awal masa berlangganan. Adapun riwayat pembayaran digunakan untuk memperoleh informasi mengenai aktivitas pembayaran pelanggan setiap bulan. Hasil seleksi data pada penelitian ini disajikan pada Tabel 1.

Tabel 1. Seleksi Data Pelanggan

ID Pelanggan	Tanggal Pemasangan	Januari	Februari	Maret	…	Desember
CM-0001	02/08/2020	cash 14/01/25	cash 08/02/25	cash 05/03/25	…	cash 11/12/25
CM-0002	09/08/2020	cash 08/01/25	cash 09/02/25	cash 10/03/25	…	cash 16/12/25

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2709
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

ID Pelanggan	Tanggal Pemasangan	Januari	Februari	Maret	…	Desember
CM-0003	05/08/2020	cash 16/01/25	cash 12/02/25	cash 10/03/25	…	cash 13/12/25
CM-0006	01/09/2020	cash 14/01/25	cash 11/02/25	cash 17/03/25	…	cash 16/12/25
CM-0008	01/09/2020	cash 03/01/25	cash 01/02/25	cash 03/03/25	…	cash 16/12/25
…	…	…	…	…	…	…
KARAN-0033	13/12/2025	-	-	-	…	man 13/12/25
BENDIL-0035	22/03/2024	bri 01/01/25	bri 04/02/25	bri 02/03/25	…	bri 04/12/25
BENDIL-0036	22/03/2024	bri 08/01/25	bri 03/02/25	bri 15/03/25	…	bri 07/12/25
BENDIL-0037	25/03/2024	bri 04/01/25	bri 09/02/25	bri 15/03/25	…	bri 02/12/25
BENDIL-0038	06/05/2024	bri 16/01/25	bri 16/02/25	bri 19/03/25	…	bri 05/12/25

Selain data operasional perusahaan, proses seleksi data juga dilakukan terhadap variabel hasil penyebaran kuesioner. Variabel yang dipilih meliputi kepuasan layanan teknisi, kepuasan layanan *customer service*, dan niat berpindah layanan. Ketiga variabel tersebut digunakan untuk merepresentasikan tingkat kepuasan pelanggan terhadap kualitas layanan yang diberikan serta kecenderungan perilaku pelanggan dalam mempertahankan atau menghentikan penggunaan layanan. Setiap variabel direpresentasikan menggunakan nilai numerik dari jawaban responden. Kategori penilaian menggunakan skala likert 1-4, pada variabel kepuasan layanan nilai 4 menunjukkan bahwa tingkat kepuasan tinggi. Namun pada variabel niat berpindah nilai semakin kecil menunjukkan bahwa adanya niatan untuk berpindah layanan. Hasil responden disajikan pada Tabel 2.

Tabel 2. Seleksi Data Kuesioner

Nama Responden	Kepuasan Layanan Teknisi	Kepuasan Layanan *Customer Service*	Niat Berpindah
Heru	4	4	4
Amri	4	4	4
Erika Dwi Septina	4	4	4
Sururi	4	4	4
Indra Cahyono	4	4	4
…	…	…	…
Mukroni	4	4	4
Juniati	4	4	3
Endang Purwati	4	4	4
Gunik	4	4	3
Mustakim	4	4	3

3.2 Pra-Pemrosesan Data

Pada tahap pra-pemrosesan data, dilakukan beberapa penyesuaian terhadap data mentah agar siap digunakan dalam proses analisis. Penyesuaian tersebut meliputi perhitungan lama berlangganan dalam satuan bulan berdasarkan tanggal pemasangan,serta perhitungan persentase keterlambatan pembayaran berdasarkan riwayat pembayaran pelanggan. Perhitungan lama berlangganan dilakukan dengan menghitung selisih antara tanggal pemasangan layanan dan periode akhir pengambilan data, kemudian dikonversi ke dalam satuan bulan. Hasil perhitungan tersebut menunjukkan perbedaan durasi berlangganan antar pelanggan, sehingga dapat digunakan untuk mengidentifikasi pelanggan yang memiliki masa berlangganan relatif singkat maupun Panjang. Pemrosesan data lama berlangganan disajikan pada Tabel 3.

Tabel 3. Hasil Prapemrosesan Lama Berlangganan

ID Pelanggan	*Start*	Pra-pemrosesan Lama Berlangganan (Bulan)
CM-0001	02/08/2020	69
CM-0002	09/08/2020	69
CM-0003	05/08/2020	69
CM-0006	01/09/2020	68
CM-0008	01/09/2020	68
…	…	…
KARAN-0033	13/12/2025	5
BENDIL-0035	22/03/2024	26
BENDIL-0036	22/03/2024	26
BENDIL-0037	25/03/2024	26
BENDIL-0038	06/05/2024	24

Pada variabel frekuensi pembayaran, dilakukan perhitungan persentase keterlambatan pembayaran untuk mengukur tingkat kedisiplinan pelanggan dalam memenuhi kewajiban pembayaran. Persentase keterlambatan diperoleh dari perbandingan antara jumlah keterlambatan pembayaran yang terlambat dan total transaksi pembayaran, kemudian dikalikan 100%. Suatu pembayaran dikategorikan terlambat apabila dilakukan setelah tanggal jatuh tempo yang telah

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2710
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

ditetapkan perusahaan, yaitu setiap tanggal 15. Hasil perhitungan persentase keterlambatan pembayaran disajikan pada Tabel 4.

Tabel 4. Hasil Pra-Pemrosesan Frekuensi Pembayaran

ID Pelanggan	Jumlah Pembayaran	Jumlah Keterlambatan	Pemrosesan Frekuensi Pembayaran (%)
CM-0001	12	1	8,33
CM-0002	12	4	33,33
CM-0003	12	2	16,67
CM-0006	12	5	41,67
CM-0008	12	3	25
…	…	…	…
KARAN-0033	1	0	0
BENDIL-0035	12	0	0
BENDIL-0036	12	0	0
BENDIL-0037	12	0	0
BENDIL-0038	12	6	50

3.3 Transformasi Data

Proses transformasi dilakukan dengan memberikan skala atau bobot pada setiap variabel. Pemberian bobot dilakukan berdasarkan kriteria yang telah ditetapkan, sehingga nilai yang diperoleh dapat merepresentasikan karakteristik pelanggan secara lebih terstuktur. Hasil transformasi pada lama berlangganan pada Tabel 5.

Tabel 5. Transformasi Lama Berlangganan

Data	Keterangan	Transformasi
1 - 24 Bulan	Rendah	1
25 - 48 Bulan	Sedang	2
≥ 49 Bulan	Tinggi	3

Berdasarkan Tabel 5. Pemberian bobot pada variabel lama berlangganan dilakukan berdasarkan durasi lama berlangganan pelanggan. Pemberian bobot ini bertujuan untuk merepresentasikan tingkat loyalitas pelanggan, di mana semakin lama masa berlangganan maka semakin tinggi bobot yang diberikan.

Tabel 6. Transformasi Frekuensi Pembayaran

Data	Keterangan	Transformasi
Keterlambatan ≥60%	Tidak Disiplin	1
Keterlambatan 40 – <60%	Kurang Disiplin	2
Keterlambatan 20 – <40%	Cukup Disiplin	3
Keterlambatan 0 – <20%	Disiplin	4

Pada Tabel 6. Nilai tertinggi diberikan kepada pelanggan dengan tingkat keterlambatan sebesar 0 hingga kurang dari 20%, yang menunjukkan bahwa pelanggan tersebut tergolong disiplin dan tepat waktu dalam memenuhi kewajiban pembayaran. Selanjutnya, pelanggan dengan tingkat keterlambatan 20% hingga kurang dari 40% diklasifikasikan sebagai cukup disiplin dengan nilai yang lebih rendah. Sementara itu, tingkat keterlambatan sebesar 40% hingga kurang dari 60% termasuk dalam kategori kurang disiplin, yang mencerminkan bahwa pelanggan cukup sering mengalami keterlambatan pembayaran. Adapun pelanggan dengan tingkat keterlambatan lebih dari atau sama dengan 60% dikategorikan sebagai tidak disiplin dan diberikan nilai terendah karena memiliki frekuensi keterlambatan yang tinggi.

Tabel 7. Transformasi Niat Berpindah

Data Awal	Transformasi Data
Niat Berpindah	1
Tetap Berlangganan	2

Pada Tabel 7, variabel niat berpindah pelanggan ditransformasikan ke dalam skala numerik 1 dan 2 untuk memudahkan proses pengolahan data. Nilai 1 menunjukkan tingkat niat berpindah yang tinggi, sedangkan nilai 2 merepresentasikan pelanggan yang memiliki kecenderungan untuk tetap berlangganan.

Tabel 8. Transformasi Kepuasan Layanan Customer Service dan Teknisi

Data Awal	Transformasi
Sangat Tidak Puas	1
Tidak Puas	2
Puas	3
Sangat Puas	4

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2711
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

Variabel kepuasan terhadap layanan teknisi dan *customer service* ditransformasikan ke dalam skala numerik 1-4 dalam pengolahan data. Nilai 1 menunjukkan tingkat kepuasan yang sangat rendah (sangat tidak puas), sedangkan nilai 4 merepresentasikan tingkat kepuasan yang sangat tinggi terhadap layanan yang diberikan.

Setelah tahap transformasi, data dinormalisai menggunakan metode *Min-Max* untuk menyamakan skala setiap variabel ke dalam rentang nilai 0-1. Data hasil normalisasi digunakan sebagai masukan pada tahap analisis menggunakan algoritma *k-means clustering* untuk membentuk segmentasi pelanggan berdasarkan karakteristik yang dimiliki. Hasil dari proses pemerataan skala dengan *min-max* ditunjukkan pada Tabel 9.

Tabel 9. Hasil Normalisai *MinMax*

ID Pelanggan	Lama Berlangganan	Frekuensi Pembayaran	Kepuasan Layanan Teknisi	Kepuasan *Customer Service*	Niat Berpindah
CM-0001	1	1	1	1	1
CM-0002	1	0,666667	0,666667	0,666667	1
CM-0003	1	1	1	0,666667	1
CM-0006	1	0,333333	1	1	1
CM-0008	1	0,666667	1	1	1
…	…	…	…	…	…
KARAN-0033	1	1	1	1	1
BENDIL-0035	0,5	1	1	0,666667	1
BENDIL-0036	0,5	1	1	1	1
BENDIL-0037	0,5	1	0,666667	1	1
BENDIL-0038	0	0,333333	0,333333	0	0

3.4 Hasil Penerapan Algoritma K-Means Clustering

Pemilihan *centroid* awal tidak dilakukan secara acak sepenuhnya, melainkan dengan mempertimbangkan jarak antar data agar posisi *centroid* tidak terlalu berdekatan. Oleh karena itu, digunakan metode *farthest point*, yaitu memilih titik-titik yang memiliki jarak paling jauh satu sama lain sebagai pusat awal klaster. Pendekatan ini dipilih karena mampu mengurangi kemungkinan tumpang tindih antar klaster serta memberikan penyebaran pusat klaster yang merata.

Tabel 10. Hasil Penentuan *Centroid* Awal

Centroid	Id Pelanggan	Lama Berlangganan	Frekuensi Pembayaran	Kepuasan Layanan Teknisi	Kepuasan *Customer Service*	Niat Berpindah
C1 (data ke-103)	GONDO-0003	0,5	1	0,666667	0,666667	1
C2 (data ke- 32)	CM-0071	0	0	0	0	0
C3 (data ke- 455)	BANGG-0002	1	1	0	0	0

Berdasarkan Tabel 10. Hasil inisialisasi centroid awal menunjukkan bahwa klaster 1 menggunakan data pelangan dengan ID GONDO-0003, klaster 2 menggunakan data pelanggan dengan ID CM-0071, dan klaster 3 menggunakan data pelanggan dengan ID BANGG-0002 sebagai titik pusat awal. Setelah *centroid* awal ditetapkan, dilakukan perhitungan jarak antara setiap data pelanggan dan masing-masing *centroid* menggunakan metode *Euclidean Distance*. Perhitungan ini bertujuan untuk mengukur tingkat kedekatan data terhadap seluruh *centroid*. Selanjutnya, setiap data ditempatkan pada klaster yang memiliki nilai jarak paling kecil (*minimum*), karena *centroid* dengan jarak terdekat dianggap paling merepresentasikan karakteristik data tersebut. Hasil pengelompokan pada iterasi pertama menjadi dasar untuk memperbarui nilai *centroid* pada iterasi berikutnya.

Tabel 11. Hasil Perhitungan Iterasi Ke-1

Id pelanggan	C1	C2	C3	Nilai Minimum	Klaster
CM-0001	0,687184	2,236068	1,732051	0,687184	1
CM-0002	0,600925	1,825742	1,414214	0,600925	1
CM-0003	0,600925	2,108185	1,563472	0,600925	1
CM-0006	0,957427	2,027588	1,855921	0,957427	1

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

Id pelanggan	C1	C2	C3	Nilai Minimum	Klaster
CM-0008	0,763763	2,108185	1,763834	0,763763	1
…	…	…	…	…	…
KARAN-0033	0,687184	2,236068	1,732051	0,687184	1
BENDIL-0035	0,333333	1,922094	1,641476	0,333333	1
BENDIL-0036	0,471405	2,061553	1,802776	0,471405	1
BENDIL-0037	0,333333	1,922094	1,641476	0,333333	1
BENDIL-0038	1,5	0,471405	1,247219	0,471405	2

Setelah proses pengelompokan pada iterasi pertama, nilai *centroid* diperbarui berdasarkan rata-rata seluruh data yang menjadi anggota pada masing-masing klaster. *Centroid* baru tersebut kemudian digunakan sebagai acuan dalam perhitungan jarak pada iterasi berikutnya. Berdasarkan hasil perhitungan pada iterasi kedua, terdapat 10 data yang mengalami perpindahan klaster. Perubahan keanggotaan tersebut menunjukkan bahwa proses pengelompokan belum mencapai kondisi stabil, sehingga algoritma masih melakukan penyesuaian terhadap posisi *centroid* untuk memperoleh struktur klaster yang lebih optimal.

Tabel 12. Data Berpindah Klaster Iterasi Ke-2

Id Pelanggan	Klaster Baru	Klaster Lama
BENDP-0061	3	1
CBC1-0054	3	2
CBC1-0064	3	2
CBC1-0070	3	2
CBC1-0079	3	1
CBC1-0097	3	2
CBK-0027	3	2
KALIP-0005	3	2
TRITI-0013	3	2
JENGG-0004	3	2

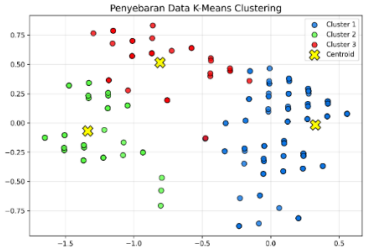
Karena pada iterasi kedua masih terdapat data yang mengalami perpindahan klaster, proses perhitungan jarak kembali dilakukan menggunakan *centroid* hasil pembaruan. Berdasarkan hasil perhitungan pada iterasi ketiga, hanya terdapat satu data yang mengalami perubahan keanggotaan klaster, yaitu pelanggan dengan ID SUMBE-0002, yang berpindah dari klaster 1 ke klaster 3. Berkurangnya jumlah data yang berpindah menunjukkan bahwa posisi *centroid* mulai stabil dan proses pengelompokan semakin mendekati kondisi optimal.

Selanjutnya, *centroid* kembali diperbarui dan digunakan dalam proses perhitungan pada iterasi keempat. Hasil perhitungan menunjukkan bahwa tidak terdapat lagi data yang mengalami perpindahan klaster. Kondisi tersebut menandakan bahwa algoritma *K-Means Clustering* telah mencapai keadaan konvergen, sehingga proses iterasi dihentikan. *Centroid* yang diperoleh pada iterasi keempat ditetapkan sebagai *centroid* akhir dan digunakan sebagai dasar dalam pembentukan klaster pelanggan. Hasil *centroid* pada iterasi keempat disajikan pada Tabel 13.

Tabel 13. Hasil Nilai *Centroid* Iterasi Ke-4

Centroid	Lama Berlangganan	Frekuensi Pembayaran	Kepuasan Layanan Teknisi	Kepuasan *Customer Service*	Niat Berpindah
C1	0,8138	0,8832	0,8694	0,8625	0,9987
C2	0,1801	0,1843	0,207	0,1946	0
C3	0,8936	0,5461	0,3546	0,2979	0,0213

Untuk memperjelas hasil pengelompokan, dilakukan visualisasi penyebaran data pada setiap klaster dalam bentuk grafik. Visualisai tersebut menunjukkan data berdasarkan kedekatannya terhadap *centroid* masing-masing klaster, sehingga pola pemisahan antar klaster dapat diamati dengan jelas. Melalui Gambar 1, proses segmentasi yang dihasilkan oleh algoritma *k-mean* mempermudahkan interpretasi terhadap karakteristik dan penyebaran anggota pada setiap klaster.



Gambar 1. Hasil Penyebaran *K-Means Clustering*

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2713
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

Pada Gambar 1. Terlihat bahwa *centroid* pada masing-masing klaster memiliki jarak yang relatif berjauhan sehingga menunjukkan pemisahan klaster yang cukup baik. Sebaran data pada setiap klaster juga cenderung mengelompok di sekitar *centroid*, yang mengindikasikan bahwa anggota klaster memiliki tingkat kemiripan karakteristik yang tinggi. Hasil segmentasi menunjukkan bahwa klaster 1 terdiri dari 776 data, klaster 2 sebanyak 161 data, dan klaster 3 sebanyak 47 data.

Selanjutnya, dilakukan proses profiling klaster untuk mengidentidikasi karakteristik setiap kelompok pelanggan berdasarkan nilai variabel penyusunnya. Proses ini bertujuan memberikan interpretasi terhadap hasil segmentasi sehingga masing-masing klaster dapat diklasifikasikan sebagai kelompok pelanggan loyal, pelanggan rentan, dan pelanggan berpotensi *churn*. Profiling dilakukan dengan mengacu pada hasil transformasi data pada setiap variabel yang telah dijelaskan pada tahap sebelumnya, sehingga karakteristik dan perilaku pelanggan pada setiap klaster dapat dianalisis secara lebih komprehensif.

Tabel 14. Hasil Profiling Penentuan Karakteristik

Klaster	Kategori	Karakteristik
1	Loyal	Pelanggan memiliki lama berlangganan tinggi, frekuensi pembayaran disiplin, kepuasan terhadap layanan teknisi sangat puas, kepuasan terhadap *customer service* sangat puas, serta tidak ada niat berpindah.
3	Rentan	Pelanggan memiliki lama berlangganan tinggi, frekuensi pembayaran cukup disiplin, kepuasan terhadap layanan teknisi tidak puas, kepuasan terhadap *customer service* tidak puas, serta memiliki niat berpindah.
2	Berpotensi *Churn*	Pelanggan memiliki lama berlangganan rendah, frekuensi pembayaran tidak disiplin, kepuasan terhadap layanan teknisi sangat tidak puas, kepuasan terhadap *customer service* sangat tidak puas, serta memiliki niat berpindah.

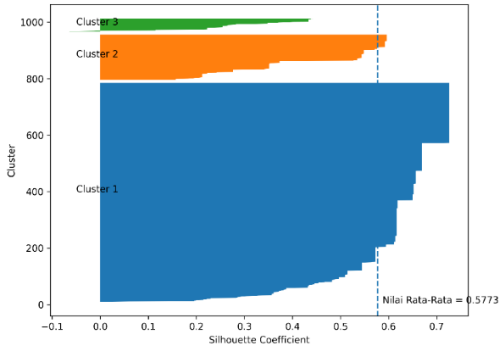
3.5 Hasil Evaluasi Silhouette Coefficient

Nilai *Silhouette Coefficient* diperoleh melalui perhitungan nilai $a(i)$, yang merepresentasikan rata-rata jarak terhadap anggota pada klaster yang sama, dan nilai $b(i)$, yaitu rata-rata jarak *minimum* data terhadap anggota pada klaster yang berbeda. Selanjutnya, kedua nilai tersebut digunakan untuk menghitung nilai *silhouette* setiap data, sehingga diperoleh ukuran kualitas keseluruhan dari hasil segmentasi yang telah terbentuk.

Tabel 15. Hasil Perhitungan Nilai *Silhouette Coefficient*

Id Pelanggan	*a(i)*	*b(i)*	*s(i)*
CM-0001	0,41908	1,529534	0,726008
CM-0002	0,559037	1,227815	0,54469
CM-0003	0,485493	1,412063	0,656181
CM-0006	0,72704	1,483059	0,50977
CM-0008	0,511826	1,47104	0,652066
…	…	…	…
KARAN-0033	0,41908	1,529534	0,726008
BENDIL-0035	0,568337	1,461706	0,611182
BENDIL-0036	0,521311	1,575703	0,669157
BENDIL-0037	0,568642	1,47416	0,61426
BENDIL-0038	0,51005	1,096778	0,534956

Hasil dari perhitungan tersebut divisualisasikan dalam bentuk grafik untuk memberikan gambaran mengenai kualitas klaster yang terbentuk. Visualisasi tersebut menunjukkan nilai *silhouette* pada setiap klaster sehingga tingkat *cohesi* dan keterpisahan antar klaster dapat diamati dengan jelas. Hasil perhitungan evaluasi klaster disajikan pada Gambar 2.



Gambar 2. Grafik *Silhouette Coefficient*

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

Berdasarkan nilai *Silhouette Coefficient* sebesar 0,5773 menunjukkan bahwa hasil pengelompokan memiliki kualitas pemisahan yang cukup baik. Nilai tersebut menggambarkan bahwa secara keseluruhan data pelanggan cenderung memiliki kedekatan yang lebih tinggi dengan anggota klasternya sendiri dibandingkan dengan anggota klaster lainnya. Meskipun demikian, nilai tersebut tidak menunjukkan bahwa seluruh anggota klaster memiliki tingkat kesesuaian yang sama. Perbedaan karakteristik dan penyebaran data pada masing-masing kelompok dapat menyebabkan variasi nilai *Silhouette* antar klaster.

Visualisasi pada Gambar 2, klaster 3 menunjukkan nilai *Silhouette* yang relatif lebih rendah dibandingkan klaster lainnya. Kondisi tersebut diduga berkaitan dengan karakteristik klaster 3 yang memiliki kombinasi atribut yang berbeda, yaitu lama berlangganan relatif tinggi, kedisiplinan pembayaran pada tingkat menengah, tetapi kepuasan terhadap layanan teknisi dan *customer service* relatif rendah serta memiliki kecenderungan berpindah layanan yang tinggi. Karakteristik tersebut memperlihatkan adanya kombinasi antara atribut yang menyerupai klaster 1 dan atribut yang mendekati klaster 2. Perbedaan pola karakteristik tersebut berpotensi menyebabkan sebagian anggota klaster 3 memiliki kedekatan dengan kelompok lain, sehingga tingkat pemisahan klasternya menjadi lebih rendah.

3.6 Pembahasan

Hasil penelitian menunjukkan bahwa metode *K-Means Clustering* mampu mengelompokkan 984 pelanggan Wi-Fi RT/RW Net ke dalam tiga klaster berdasarkan karakteristik lama berlangganan, frekuensi pembayaran, kepuasan terhadap layanan teknisi, kepuasan terhadap *customer service*, dan niat berpindah layanan. Berdasarkan proses profiling, klaster 2 diidentifikasi sebagai kelompok pelanggan yang berpotensi *churn* karena memiliki karakteristik berupa lama berlangganan yang relatif rendah, tingkat keterlambatan pembayaran yang tinggi, kepuasan layanan yang sangat rendah, serta kecenderungan untuk berpindah layanan. Sebaliknya, klaster 1 didominasi oleh pelanggan yang loyal dengan pembayaran disiplin dan tingkat kepuasan yang tinggi, sedangkan klaster 3 menunjukkan kondisi pelanggan yang mulai berisiko karena memiliki kepuasan layanan yang rendah meskipun telah berlangganan dalam waktu yang cukup lama.

Berdasarkan karakteristik masing-masing klaster, terdapat perbedaan kondisi pelanggan yang dapat digunakan sebagai dasar dalam menentukan prioritas pengelolaan layanan. Klaster 1 merupakan kelompok pelanggan loyal dengan jumlah anggota terbesar, yaitu 776 pelanggan atau 78,86% dari keseluruhan data. Tingginya nilai pada variabel lama berlangganan, kedisiplinan pembayaran, serta kepuasan terhadap layanan teknisi dan *customer service* menunjukkan bahwa kelompok ini memiliki hubungan yang relatif baik dengan perusahaan.

Berbeda dengan klaster 1, klaster 2 terdiri atas 161 pelanggan atau 16,36% yang menunjukkan karakteristik paling mengkhawatirkan. Kelompok ini memiliki masa berlangganan relatif rendah, kedisiplinan pembayaran yang rendah, serta tingkat kepuasan terhadap layanan teknisi dan *customer service* yang rendah. Selain itu, nilai variabel niat berpindah menunjukkan kecenderungan berpindah layanan yang tinggi. Kombinasi karakteristik tersebut mengindikasikan bahwa pelanggan dalam klaster 2 perlu mendapatkan perhatian lebih awal karena berpotensi menghentikan penggunaan layanan.

Sementara itu, klaster 3 terdiri atas 47 pelanggan atau 4,78% dengan karakteristik yang berbeda dari kedua klaster lainnya. Kelompok ini memiliki masa berlangganan relatif tinggi, tetapi tingkat kepuasan terhadap layanan teknisi dan *customer service* tergolong rendah. Kedisiplinan pembayaran juga berada pada tingkat menengah. Kondisi tersebut menunjukkan bahwa lamanya pelanggan menggunakan layanan tidak selalu mencerminkan tingkat kepuasan yang tinggi.

Berdasarkan hasil segmentasi tersebut, perusahaan dapat menerapkan strategi retensi pelanggan yang disesuaikan dengan karakteristik setiap klaster. Pada klaster 1, strategi yang dapat dilakukan adalah mempertahankan loyalitas pelanggan melalui peningkatan konsistensi kualitas jaringan, pemeliharaan infrastruktur secara berkala, serta pemberian apresiasi kepada pelanggan yang memiliki riwayat berlangganan panjang dan pembayaran disiplin. Pada klaster 2, perusahaan perlu memprioritaskan pendekatan secara langsung melalui komunikasi dengan pelanggan untuk mengidentifikasi penyebab ketidakpuasan, melakukan pemeriksaan kualitas jaringan, mempercepat penanganan gangguan, serta meningkatkan responsivitas *customer service*. Perusahaan juga dapat memberikan pengingat pembayaran dan mengevaluasi kesesuaian paket layanan dengan kebutuhan pelanggan. Sementara itu, penanganan Klaster 3 dapat difokuskan pada upaya pencegahan perpindahan pelanggan melalui pemeriksaan jaringan secara proaktif, evaluasi riwayat keluhan, tindak lanjut terhadap ketidakpuasan layanan teknisi dan *customer service*, serta komunikasi berkala untuk mengetahui kebutuhan pelanggan. Strategi tersebut diharapkan dapat membantu perusahaan menentukan prioritas pelayanan dan mengoptimalkan upaya mempertahankan pelanggan berdasarkan karakteristik masing-masing kelompok.

Karakteristik pada klaster 2 sejalan dengan penelitian yang dilakukan Hakim (2024) [19], [20] yang menyatakan bahwa rendahnya kualitas layanan dan kepuasan pelanggan merupakan faktor utama penyebab *churn*. Selain itu, Alfiansyah dkk. (2025) menjelaskan bahwa pelanggan dengan riwayat pembayaran yang tidak disiplin memiliki risiko churn yang lebih tinggi, sedangkan Rahmadiana dan Lestarini (2025) menunjukkan bahwa pelanggan cenderung berpindah layanan ketika pengalaman layanan yang diterima tidak sesuai dengan harapan. Dengan demikian, kombinasi karakteristik pada klaster 2 menunjukkan rendahnya loyalitas pelanggan dan tingginya potensi untuk menghentikan penggunaan layanan.

Hasil penelitian ini juga memperkuat konsep dasar metode *K-Means* dalam mengelompokkan data berdasarkan kemiripan karakteristik. Temuan ini konsisten dengan penelitian Harani dkk. (2020) dan Rahmadiana (2025) yang menunjukkan bahwa *K-Means* mampu menghasilkan segmentasi pelanggan dengan tingkat risiko *churn* yang berbeda sehingga dapat mendukung penyusunan strategi retensi pelanggan. Evaluasi menggunakan *Silhouette Coefficient* menghasilkan nilai 0,5773, yang termasuk dalam kategori baik. Nilai tersebut menunjukkan bahwa klaster yang terbentuk memiliki tingkat kedekatan data dalam klaster yang tinggi dan pemisahan antar klaster yang cukup jelas [21].

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2715
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2705-2718
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1248

4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan metode *K-Means Clustering* untuk melakukan segmentasi terhadap pelanggan 984 Wi-Fi RT/RW Net dalam mengidentifikasi potensi *customer churn*. Proses segmentasi dilakukan menggunakan variabel lama berlangganan, frekuensi pembayaran, kepuasan terhadap layanan teknisi, kepuasan terhadap layanan *customer service*, dan niat berpindah pelanggan. Melalui proses pengelompokan secara iteratif berdasarkan kedekatan jarak terhadap *centroid*, diperoleh tiga klaster dengan karakteristik yang berbeda, yaitu klaster 1 sebanyak 776 pelanggan, klaster 2 sebanyak 161 pelanggan, dan klaster 3 sebanyak 47 pelanggan. Berdasarkan hasil *profiling*, klaster 1 merupakan kelompok pelanggan loyal, klaster 2 merupakan kelompok pelanggan berpotensi *churn*, sedangkan klaster 3 merupakan kelompok pelanggan rentan. Berdasrkan karakteristik tersebut, perusahaan perlu menerapkan strategi retensi yang berbeda pada setiap klaster. Klaster 1 perlu dipertahankan melalui peningkatan kualitas jaringan, pelayanan, dan program loyalitas. Klaster 2 perlu menjadi prioritas penanganan melalui komunikasi langsung, penyelesaian keluhan, dan penawaran layanan sesuai kebutuhan pelanggan. Sementara itu, klaster 3 memerlukan pemantauan kualitas jaringan, evaluasi pelayanan teknisi dan *customer service*, serta komunikasi proaktif untuk mengidentifikasi ketidakpuasan pelanggan. Strategi tersebut diharapkan dapat membantu CV. Bagus Agung Mandiri Sejahtera dalam menentukan prioritas pengelolaan pelanggan dan mengurangi potensi *churn*. Hasil evaluasi menggunakan *Silhouette Coefficient* memperoleh nilai sebesar 0,5773, yang menunjukkan bahwa struktur klaster yang terbentuk memiliki tingkat kedekatan data dalam klaster yang baik serta pemisahan antar klaster yang cukup jelas. Dengan demikian, metode *K-Means Clustering* mampu menghasilkan segmentasi pelanggan yang dapat dimanfaatkan sebagai dasar dalam penyusunan strategi retensi pelanggan, khususnya untuk mengidentifikasi pelanggan loyal, pelanggan yang mulai rentan, dan pelanggan yang berpotensi *churn* dalam layanan RT/RW Net. Meskipun demikian, penelitian ini masih memiliki keterbatasan karena hanya menggunakan lima variabel dan menerapkan metode *K-Means* dengan *Euclidean Distance*, jumlah klaster yang telah ditentukan sesuai dengan kebutuhan penelitian, serta evaluasi menggunakan *Silhouette Coefficient*. Oleh karena itu, penelitian selanjutnya disarankan untuk menambahkan variabel lain yang berpotensi memengaruhi *churn*, seperti intensitas penggunaan layanan, riwayat komplain, dan nominal pembayaran bulanan, serta membandingkan metode penentuan jumlah klaster, teknik pengukuran jarak, maupun metode evaluasi lainnya agar diperoleh hasil segmentasi yang lebih optimal dan akurat.

5. PERNYATAAN PENULIS

5.1 Kontribusi Penulis

Konseptualisasi: M. D. H., H. Y., dan W. D. P.;
Metodologi: M. D. H., H. Y., dan W. D. P.;
Perangkat Lunak: M. D. H.;
Validasi: H. Y., dan W. D. P.;
Analisis Formal: M. D. H., H. Y., dan W. D. P.;
Investigasi: M. D. H., H. Y., dan W. D. P.;
Sumber Daya: M. D. H., H. Y., dan W. D. P.;
Kurasi Data: M. D. H.;
Penulisan Draf Awal: M. D. H.;
Penulisan, Peninjauan, dan Penyuntingan: M. D. H., H. Y., dan W. D. P.;
Visualisasi: M. D. H., H. Y., dan W. D. P.;
Seluruh penulis telah membaca dan menyetujui versi manuskrip yang telah diterbitkan.

5.2 Ketersediaan Data

Data yang disajikan dalam penelitian ini tersedia berdasarkan permintaan kepada penulis korespondensi.

5.3 Pendanaan

Penelitian ini tidak memperoleh pendanaan dari lembaga pemerintah, swasta, maupun pihak eksternal lainnya. Seluruh proses penelitian dilaksanakan secara mandiri tanpa dukungan pendanaan eksternal.

5.4 Pernyataan Dewan Peninjau Institusional

Penelitian ini dilaksanakan berdasarkan izin dari pihak perusahaan terkait untuk memperoleh dan menggunakan data yang diperlukan dalam penelitian. Pengumpulan data dilakukan melalui penyebaran kuesioner kepada pelanggan serta pengambilan data perusahaan dengan tetap memperhatikan kerahasiaan informasi dan privasi responden. Penelitian ini tidak memperoleh persetujuan khusus dari Dewan Peninjau Institusional (Institutional Review Board).

5.5 Pernyataan Persetujuan Berdasarkan Informasi

Persetujuan partisipasi dalam penelitian ini diperoleh dari setiap responden sebelum pengisian kuesioner dilakukan. Responden terlebih dahulu diberikan penjelasan mengenai tujuan dan kepentingan penelitian serta berpartisipasi secara

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

sukarela. Seluruh data dan informasi yang diperoleh digunakan hanya untuk kepentingan penelitian dengan tetap menjaga kerahasiaan identitas dan privasi responden.

5.6 Pernyataan Konflik Kepentingan

Para penulis menyatakan bahwa mereka tidak memiliki kepentingan keuangan yang bersaing atau hubungan pribadi yang diketahui yang dapat dianggap memengaruhi pekerjaan yang dilaporkan dalam artikel ini.

5.7 Penggunaan AI Generatif dan Teknologi Berbantuan AI dalam Proses Penulisan

Para penulis menggunakan AI Generatif untuk meningkatkan kejelasan penulisan dalam artikel ini. Para penulis telah meninjau dan menyunting konten yang dihasilkan dengan bantuan AI serta bertanggung jawab penuh atas versi akhir publikasi.

REFERENCES

[1] A. P. J. I. Indonesia, "Survei Penetrasi Internet dan Perilaku Penggunaan Internet." Diakses: 12 November 2025. [Daring]. Tersedia pada: https://survei.apjii.or.id/

[2] R. N. Fathimatuzzahra dan T. Dompak, "Menjembatani Kesenjangan Akses Internet : Studi Kasus Indonesia dan Brunei Darussalam," *Pros. Semin. Nas. Ilmu Sosisal Teknol.*, vol. 6, hal. 147–153, 2024, doi: https://doi.org/10.33884/psnistek.v6i.

[3] F. I. Maulana, F. N. Safitri, A. M. Samhan, dan B. E. Tomo, "Peluang Usaha RT/RW Net," *Semin. Nas. Call Pap. HUBISINTEK 2023*, vol. Vol. 4 No., hal. 284–294, 2023, [Daring]. Tersedia pada: https://www.ojs.udb.ac.id/HUBISINTEK/article/view/3612/2436

[4] R. R. Alfiansyah, E. Purwanto, dan D. R. Irawan, "Segmentasi Pelanggan Jaringan Wifi RT/RW Net Menggunakan Metode K-Means Clustering Untuk Analisis Churn (Studi Kasus: PT Nusa Data Multimedia)," *Semin. Nas. Teknol. Inf. dan Bisnis*, hal. 349–353, 2025, doi: 10.47701/38s0pt05.

[5] F. Naufal, Martanto, A. R. Dikananda, dan D. Rohman, "Analisis Segmentasi Pelanggan Voucher Wifi Dengan Metode K-Means," *J. Inform. Teknol. dan Sains*, vol. 7, no. 1, hal. 81–90, 2025, doi: https://doi.org/10.51401/jinteks.v7i1.5169.

[6] N. Sukarno Wijaya, M. Jajuli, dan B. Arif Dermawan, "Penerapan Algoritma K-Means Clustering Dalam Menentukan Daerah Prioritas Penanganan Kemiskinan Di Wilayah Jawa Timur," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 4, hal. 7579–7584, 2024, doi: 10.36040/jati.v8i4.10248.

[7] N. H. Ahsina, F. Fatimah, dan F. Rachmawati, "Analisis Segmentasi Pelanggan Bank Berdasarkan Pengambilan Kredit Dengan Menggunakan Metode K-Means Clustering," *J. Ilm. Teknol. Inf. Terap.*, vol. 8, no. 3, hal. 393–400, 2022, doi: https://doi.org/10.33197/jitter.vol8.iss3.2022.883.

[8] A. A. D. Sulistyawati dan M. Sadikin, "Penerapan Algoritma K - Medoids untuk Menentukan Segmentasi Pelanggan," *Sist. J. Sist. Inf.*, vol. 10, hal. 516–526, 2021, [Daring]. Tersedia pada: https://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/view/1332/394

[9] N. D. Azahra, Z. L. Maulida, M. F. Farizqi, dan I. Kharisudin, "Segmentasi Pelanggan dengan Algoritma DBSCAN dan K-Means : Mengungkap Pola dalam Analisis Kepribadian Pelanggan," *Pros. Semin. Nas. Mat.*, vol. 8, hal. 370–376, 2025, [Daring]. Tersedia pada: https://proceeding.unnes.ac.id/prisma/article/view/4337/3860

[10] A. M. Artiarno, P. Setiaji, dan F. Nugraha, "K-Means Clustering untuk Segmentasi Pelanggan: Mengungkap Pola Pembelian Strategi Pemasaran pada Sektor Ritel," *Edumatic J. Pendidik. Inform.*, vol. 9, no. 2, hal. 442–451, 2025, doi: 10.29408/edumatic.v9i2.30336.

[11] W. T. Oktaviany, F. Insani, A. Nazir, dan Pizaini, "Pengelompokan Wilayah Bencana Banjir di Indonesia Menggunakan Algoritma K-Means," *Bull. Comput. Sci. Res.*, vol. 5, no. 4, hal. 542–553, 2025, doi: 10.47065/bulletincsr.v5i4.608.

[12] A. Julian, A. Faqih, dan G. Dwilestari, "Segmentasi Pelanggan Menggunakan Algoritma K-Means Pada Jaringan Telekomunikasi Untuk Optimalisasi Strategi Pemasaran," *JTT (Jurnal Teknol. Terpadu)*, vol. 13, no. 1, hal. 94–100, 2025, doi: 10.32487/jtt.v13i1.2436.

[13] R. N. Rahmadiana dan D. Lestarini, "Implementasi Algoritma K-Means untuk Pengelompokan Customer Churn dalam Menentukan Strategi Pemasaran Bank," *Sist. J. Sist. Inf.*, vol. 14, no. 4, hal. 1909–1919, 2025, [Daring]. Tersedia pada: http://sistemasi.ftik.unisi.ac.id

[14] Y. Gunawan, "Analisis Faktor-Faktor Pengaruh Customer Churn Pada Sebuah Perusahaan Internet Provider," *J. Konseling dan Pendidik.*, vol. 14, no. 1, hal. 577–590, 2026, [Daring]. Tersedia pada: https://jurnal.konselingindonesia.com/index.php/jkp/about/submissions#authorGuidelines

[15] N. H. Harani, C. Prianto, dan F. A. Nugraha, "Segmentasi Pelanggan Produk Digital Service Indihome Menggunakan Algoritma K-Means Berbasis Python," *J. Manaj. Inform.*, vol. 10, no. 2, hal. 133–146, 2020, doi: 10.34010/jamika.v10i2.2683.

[16] N. Salma dan A. Aprianingsih, "Customer Churn Analysis : Analyzing Customer Churn Determinants on an ISP Company in Indonesia," vol. 19, no. 1, hal. 29–40, 2021, doi: 10.17933/bpostel.2021.190103.

[17] P. I. Pangestu, T. I. Hermanto, dan D. Irmayanti, "Analisis Segmentasi Pelanggan Berbasis Model Recency Frequency Dan Monetary (RFM) Menggunakan Algoritma K-Means," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3s1, 2023, doi: 10.23960/jitet.v11i3s1.3396.

[18] S. A. Perdana, S. F. Florentin, dan A. Santoso, "Analisisi Segmentasi Pelanggan Menggunakan K-Means Clustering Studi Kasus Aplikasi AlfaGift," *Sebatik*, vol. 26, no. 2, hal. 446–457, 2022, doi: 10.46984/sebatik.v26i2.2134.

[19] R. N. S. Hakim dan Asmunin, "Optimasi Algoritma Random Forest dengan Teknik Boosting dalam Prediksi Churn Pelanggan di Industri Telekomunikasi," *J. Manag. Inform.*, hal. 1–9, 2024, [Daring]. Tersedia pada: https://ejournal.unesa.ac.id/index.php/jurnal-manajemen-informatika/article/view/59841

[20] A. R. Y. Siregar dan M. Iqbal, "Prediksi Customer Churn Pada Layanan Indihome Menggunakan Algoritma Decision Tree (Studi Kasus Pt. Telkom Akses)," *J. Sci. Soc. Res.*, vol. 8, no. 1, hal. 204–211, 2025, [Daring]. Tersedia pada:

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2717
This work is licensed under a Creative Commons Attribution 4.0 International License.

https://jurnal.goretanpena.com/index.php/JSSR/article/view/2698/1512

[21] F. D. S. Alhamdani, A. A. Dianti, dan Y. Azhar, "Segmentasi Pelanggan Berdasarkan Perilaku Penggunaan Kartu Kredit Menggunakan Metode K-Means Clustering," *J. Inform. Sunan Kalijaga*, vol. 6, no. 2, hal. 70–77, 2021, doi: https://doi.org/10.14421/jiska.2021.6.2.70-77.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2718
This work is licensed under a Creative Commons Attribution 4.0 International License.