

Strategi Customer Intelligence Melalui Klasterisasi K-Means untuk Optimalisasi Retensi Pelanggan pada Industri Telekomunikasi

Muhammad Azril Surya Ramadhan^{1*}, Winda Chariska², Mufidah Karimah³

^{1,2,3}Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Pamulang, Tangerang Selatan, Indonesia

¹azrilramadhan949@gmail.com, ²windacariska913@gmail.com, ³dosen02829@unpam.ac.id
(Email Correspondence: azrilramadhan949@gmail.com)

Received: 4 April 2026 | Revised: 8 April 2026 | Accepted: 30 April 2026

Abstrak

Di tengah persaingan ketat industri telekomunikasi, tingginya churn rate pelanggan mengancam stabilitas pendapatan perusahaan. Penelitian ini menerapkan klasterisasi K-Means pada dataset Telco Customer Churn untuk mengidentifikasi segmen pelanggan guna optimalisasi retensi melalui customer intelligence, mengikuti kerangka CRISP-DM. Data diproses dengan imputasi median pada Total Charges, normalisasi StandardScaler, dan penentuan klaster optimal ($k=4$) via Elbow Method. Hasil mengungkap empat profil: (1) The Newbies (tenure rendah, biaya rendah); (2) High-Value Loyalists (tenure tinggi, biaya tinggi); (3) At-Risk Big Spenders (tenure rendah, biaya sangat tinggi); dan (4) Budget Veterans (tenure tinggi, biaya rendah), divalidasi Silhouette Score solid. Temuan memungkinkan strategi personalisasi: promo onboarding untuk Newbies, VIP rewards untuk Loyalists, kontrak jangka panjang untuk At-Risk, dan upselling bundling untuk Veterans. Pendekatan ini meningkatkan efisiensi alokasi sumber daya pemasaran berbasis data.

Kata kunci: Customer Intelligence, K-Means, Strategi Bisnis, Retensi Pelanggan, Telco Churn.

Abstract

Amid fierce competition in the telecommunications industry, high customer churn rates threaten the stability of a company's revenue. This study applies K-Means clustering to the Telco Customer Churn dataset to identify customer segments for retention optimization through customer intelligence, following the CRISP-DM framework. The data is processed using median imputation for Total Charges, StandardScaler normalization, and determination of the optimal number of clusters ($k=4$) via the Elbow Method. The results revealed four profiles: (1) The Newbies (short tenure, low costs); (2) High-Value Loyalists (long tenure, high costs); (3) At-Risk Big Spenders (short tenure, very high costs); and (4) Budget Veterans (long tenure, low costs), validated by a solid Silhouette Score. These findings enable personalized strategies: onboarding promotions for new customers, VIP rewards for loyal customers, long-term contracts for At-Risk customers, and upselling bundles for long-time customers. This approach improves the efficiency of data-driven marketing resource allocation.

Keywords: Customer Intelligence, K-Means, Business Strategy, Customer Retention, Telco Churn..

1. PENDAHULUAN

Dinamika industri telekomunikasi global saat ini telah mengalami pergeseran paradigma yang sangat fundamental, bergerak dari fase perluasan jaringan infrastruktur secara masif menuju persaingan yang sangat ketat dalam mempertahankan basis pengguna yang sudah ada. Penetrasi pasar yang semakin jenuh membuat pelanggan memiliki daya tawar (bargaining power) yang jauh lebih tinggi dibandingkan dekade sebelumnya. Dalam ekosistem digital yang hiper-terhubung ini, konsumen dapat dengan sangat mudah membandingkan layanan dan berpindah ke penyedia layanan kompetitor yang menawarkan insentif harga, paket bundling yang lebih menarik, atau kualitas jaringan yang dinilai lebih stabil dan

menguntungkan bagi kebutuhan spesifik mereka. Fenomena kemudahan berpindah ini tidak dapat dihindari dan menciptakan tantangan krusial bagi operator seluler berupa tingginya angka perpindahan pelanggan atau yang secara industri dikenal sebagai churn rate. Bayang-bayang fluktuasi churn rate ini terus mengancam stabilitas profitabilitas penyedia layanan. Tingginya angka churn bukan sekadar masalah hilangnya satu atau dua pengguna dari database, melainkan sebuah ancaman langsung dan terukur terhadap pendapatan inti perusahaan serta pangsa pasar secara keseluruhan.

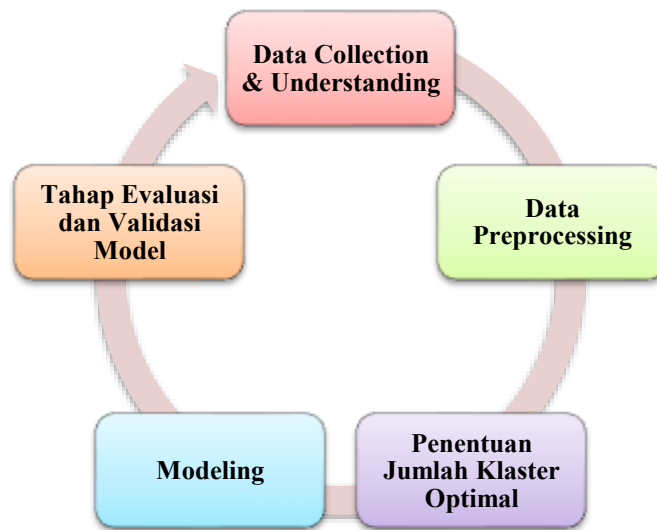
Secara finansial dan operasional bisnis, biaya yang harus dikeluarkan perusahaan untuk mengakuisisi pelanggan baru (Customer Acquisition Cost) secara historis selalu jauh lebih membebani anggaran—bahkan bisa mencapai lima hingga dua puluh lima kali lipat lebih mahal—dibandingkan dengan upaya mempertahankan pelanggan lama (Customer Retention Cost) [1]. Meskipun fakta ekonomi ini telah disadari secara luas, banyak perusahaan telekomunikasi masih terjebak dan mengandalkan strategi pemasaran serta program retensi yang bersifat usang, statis, dan terlalu umum (one-size-fits-all). Pendekatan semacam ini lambat laun kehilangan relevansi dan efektivitasnya karena secara inheren gagal mengakomodasi keragaman profil pengguna, tingkat pengeluaran finansial, serta pola interaksi masing-masing pelanggan terhadap ragam layanan yang ditawarkan. Mempertahankan pelanggan di era modern menuntut pemahaman mendalam tentang siklus hidup konsumen [9]. Tanpa adanya personalisasi, promosi yang dikirimkan seringkali diabaikan karena tidak tepat sasaran, yang pada akhirnya justru memicu ketidakpuasan pelanggan [13].

Menghadapi kompleksitas persoalan tersebut, perusahaan telekomunikasi modern memiliki urgensi absolut untuk beralih pada pendekatan pengambilan keputusan berbasis data (data-driven decision making) melalui penerapan customer intelligence yang komprehensif. Secara operasional, industri telekomunikasi sebenarnya sangat diuntungkan oleh ketersediaan volume data transaksi, log panggilan, penggunaan internet, dan rekam jejak aktivitas pelanggan yang sangat masif yang terus dihasilkan setiap harinya [11]. Kumpulan data mentah raksasa (Big Data) ini merupakan aset strategis yang tak ternilai harganya jika diekstraksi dan dianalisis secara tepat menggunakan teknologi komputasi modern. Melalui penerapan teknik machine learning dan data mining, informasi operasional harian tersebut dapat diolah untuk mengungkap pola perilaku tersembunyi (hidden patterns) yang mendasari keputusan psikologis dan finansial seorang pelanggan untuk tetap menetap atau justru berhenti berlangganan [2] [5].

Salah satu metode analitik yang paling dominan dan relevan untuk membedah keragaman perilaku konsumen yang kompleks ini adalah model klasterisasi (clustering). Algoritma klasterisasi, khususnya yang berbasis unsupervised learning, beroperasi dengan membagi basis pelanggan ke dalam segmen-segmen yang homogen berdasarkan kemiripan karakteristik multivariatnya [15]. Dengan pemetaan ini, perusahaan dapat melihat basis penggunaannya secara lebih tajam dan terstruktur. Dalam literatur dan implementasi industri, algoritma K-Means terbukti menjadi salah satu pendekatan yang paling tangguh, efisien secara komputasi, dan efektif dalam mempartisi kumpulan data pelanggan berskala besar [3] [6]. Segmentasi yang dihasilkan oleh K-Means menjadi fondasi arsitektural yang penting untuk mengenali berbagai profil di lapangan secara objektif, mulai dari kelompok pelanggan baru yang masih membutuhkan penyesuaian layanan, pelanggan setia bernilai tinggi yang menyumbang mayoritas margin, hingga mereka yang secara metrik berada di ambang keputusan kritis untuk melakukan churn [4] [12].

Berangkat dari urgensi industri dan potensi teknologi analitik tersebut, penelitian ini berfokus pada pemanfaatan teknik segmentasi pelanggan melalui metode K-Means untuk merancang strategi bisnis telekomunikasi yang jauh lebih adaptif dan proaktif [14]. Pendekatan berbasis tree-classifiers maupun klasterisasi fitur numerik terbukti mampu mengoptimalkan parameter retensi [7]. Identifikasi segmen pelanggan yang presisi dan beralasan matematis akan memungkinkan jajaran manajemen strategis untuk mengalokasikan sumber daya pemasaran, anggaran promosi, dan layanan pelanggan secara lebih rasional, tertarget, dan efisien [8] [10]. Melalui kerangka kerja analitik ini, intervensi promosi atau penawaran khusus dapat diprioritaskan secara personal khusus pada kelompok dengan risiko churn tertinggi guna mencegah perpindahan yang merugikan. Pada akhirnya, transformasi sekumpulan data mentah menjadi wawasan operasional cerdas (customer intelligence) ini diharapkan mampu menjadi solusi holistik bagi perusahaan telekomunikasi untuk tidak hanya mencegah churn, tetapi juga mengoptimalkan tingkat loyalitas, mempertahankan dominasi pangsa pasar, serta meningkatkan Customer Lifetime Value (CLV) secara eksponensial dalam orientasi bisnis jangka panjang.

2. METODE PENELITIAN



Gambar 1. Kerangka CRISP-DM

Penelitian ini dilaksanakan melalui serangkaian tahapan komputasional dan analitik yang sistematis, mengadopsi standar metodologi data sains guna memastikan keabsahan, keandalan, dan reproduksibilitas proses penggalian data. Adapun tahapan-tahapan penelitian yang dilakukan adalah sebagai berikut:

2.1 Tahap Pengumpulan dan Pemahaman Data (Data Collection & Understanding)

Tahap pertama dalam kerangka kerja difokuskan pada pengumpulan serta pemahaman mendalam terhadap struktur data mentah yang akan dianalisis. Sumber data utama yang dieksplorasi secara ekstensif adalah dataset publik Telco Customer Churn, yang merekam jejak interaksi komersial konsumen dengan penyedia layanan telekomunikasi. Pada fase krusial ini, proses eksplorasi data secara komprehensif atau Exploratory Data Analysis (EDA) dilakukan untuk menyeleksi fitur-fitur yang paling merepresentasikan nilai ekonomi pelanggan. Analisis korelasi awal diterapkan untuk menyaring atribut yang memiliki dampak signifikan terhadap perilaku retensi, sehingga dimensi data menjadi lebih efisien untuk diproses. Melalui proses seleksi heuristik tersebut, tiga variabel prediktor numerik utama berhasil diisolasi sebagai fokus utama komputasi klusterisasi. Ketiga variabel tersebut meliputi tenure (durasi bulan masa berlangganan), Monthly Charges (kewajiban tagihan rutin bulanan), serta Total Charges (akumulasi historis keseluruhan biaya yang telah dikeluarkan). Pemilihan ketiga variabel spesifik ini didasarkan pada rasionalitas bisnis bahwa dimensi metrik tersebut secara empiris mencerminkan riwayat transaksi secara langsung. Pemahaman yang kuat terhadap ketiga atribut dasar ini menjadi fondasi awal yang valid untuk mengukur tingkat komitmen pengguna terhadap ekosistem layanan perusahaan.

2.2 Tahap Pra-pemrosesan Data (Data Preprocessing)

Setelah karakteristik mentah dipahami dengan baik, tahapan krusial selanjutnya adalah pra-pemrosesan untuk mentransformasi data mentah menjadi representasi kerangka kerja yang bersih. Proses pembersihan awal diawali dengan audit kualitas data, di mana ditemukan anomali berupa nilai kosong (missing values) spesifik pada kolom Total Charges. Ketidaklengkapan informasi ini diselesaikan secara komputasional menggunakan teknik imputasi sentral yang memanfaatkan nilai tengah atau median dari distribusi data keseluruhan. Penggunaan metrik median terbukti jauh lebih superior dibandingkan rata-rata (mean) karena sifatnya yang sangat tangguh (robust) terhadap bias yang dipicu oleh keberadaan data pencilan (outliers). Selanjutnya, tahapan pra-pemrosesan mewajibkan adanya standarisasi skala karena algoritma K-Means secara inheren menggunakan perhitungan jarak geometris (Euclidean distance) dalam operasinya. Tanpa standarisasi, variabel bervolume numerik besar seperti Total Charges berpotensi

mendominasi metrik jarak dan secara otomatis menegaskan signifikansi variabel berskala kecil seperti tenure. Untuk mengeliminasi bias dimensi spasial tersebut, seluruh fitur numerik ditransformasikan secara merata menggunakan fungsi StandardScaler dari pustaka pemrosesan data. Transformasi matematis ini menghasilkan ekuivalensi skala data terpusat, sehingga setiap atribut memberikan bobot kontribusi yang proporsional pada perhitungan pusat kluster.

2.3 Tahap Penentuan Jumlah Kluster Optimal

Memasuki fase arsitektural model, salah satu tantangan sekaligus kewajiban inheren dari algoritma K-Means adalah keharusan untuk mendefinisikan hiperparameter jumlah kluster (k) secara apriori. Mengingat algoritma ini tidak dapat menentukan jumlah segmen terbaiknya sendiri, pengujian komputasi harus dilakukan untuk menemukan titik keseimbangan pemisahan data. Oleh karena itu, simulasi pemodelan dijalankan secara iteratif dengan menguji rentang nilai (k) mulai dari percobaan dua kelompok hingga maksimal sepuluh kluster. Kualitas kepadatan dan pemisahan data pada setiap iterasi dievaluasi secara komprehensif menggunakan pendekatan Elbow Method yang sangat populer dalam literatur sains data. Metode ini bekerja dengan menghitung dan memplot metrik Inertia atau yang dikenal sebagai Within-Cluster Sum of Squares (WCSS) dari masing-masing iterasi percobaan. Observasi visual dilakukan terhadap pergerakan grafik untuk mencari pola penurunan variansi intrakluster yang sangat tajam pada fase-fase awal pembentukan. Model secara spesifik mencari titik di mana kurva mulai melandai secara drastis hingga membentuk sudut bengkok yang menyerupai lipatan siku lengan manusia. Titik pelandaian atau elbow inilah yang diinterpretasikan sebagai batas arsitektur kluster yang paling optimal, di mana penambahan parameter lebih lanjut dinilai sudah tidak lagi efisien.

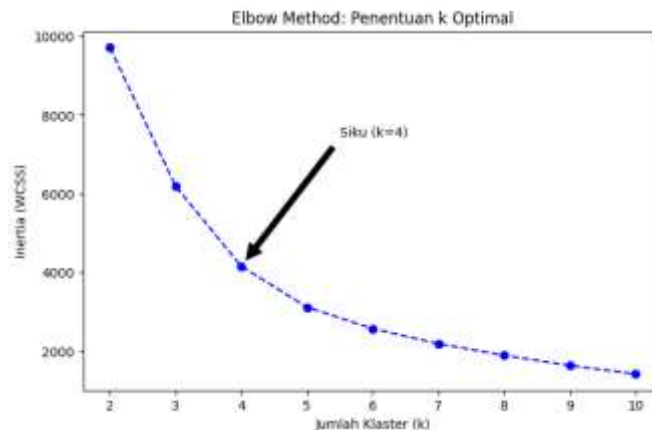
2.4 Tahap Pemodelan K-Means (Modeling)

Setelah tahapan pembersihan data dan pencarian hiperparameter jumlah kluster terselesaikan, algoritma K-Means secara formal diimplementasikan ke dalam sistem. Model difungsikan untuk melakukan pemetaan dan pemisahan data secara unsupervised, yang berarti proses pelatihan mesin berjalan mandiri tanpa menggunakan label target referensi. Proses komputasi dimulai ketika algoritma secara acak menempatkan titik-titik pusat referensi atau centroid awal sejumlah k (empat) pada ruang dimensi data yang telah dinormalisasi. Pada langkah selanjutnya, algoritma secara matematis mengkalkulasi jarak spasial antara setiap titik data pelanggan terhadap seluruh centroid yang tersedia di ruang komputasi. Masing-masing entitas pelanggan kemudian dialokasikan secara eksklusif ke dalam kelompok yang memiliki jarak Euclidean paling dekat dengan posisi centroid-nya. Setelah seluruh pelanggan terdistribusi penuh, sistem melakukan penghitungan ulang terhadap posisi rata-rata dari seluruh anggota dalam satu kluster untuk menentukan koordinat centroid yang baru. Algoritma K-Means akan terus mengulangi seluruh proses alokasi dan pembaruan koordinat pusat ini secara iteratif hingga mencapai titik konvergensi yang absolut. Titik konvergensi komputasional ini tercapai ketika posisi centroid tidak lagi mengalami pergeseran yang berarti, menandakan bahwa segmentasi profil transaksi pelanggan telah terbentuk secara final dan stabil.

2.5 Tahap Evaluasi dan Validasi Model

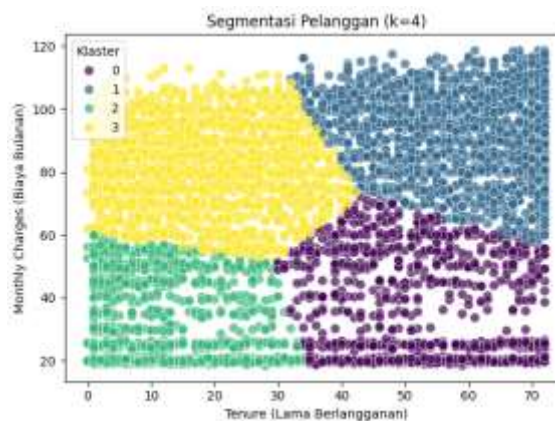
Tahap purna-pemodelan difokuskan pada proses evaluasi dan validasi matematis untuk mengukur seberapa presisi pemisahan objek yang telah dilakukan oleh arsitektur mesin. Mengingat ketiadaan data latih berlabel (ground truth) pada pendekatan unsupervised learning, evaluasi kualitas kluster tidak dapat bersandar secara subjektif pada intuisi visual semata. Untuk mengatasi keterbatasan teknis tersebut, penelitian ini menggunakan formula metrik Silhouette Score sebagai standar kuantifikasi yang sangat objektif terhadap ketegasan struktur kluster. Metrik siluet ini bekerja dari dalam dengan mengkalkulasi tingkat kohesi, yakni parameter seberapa dekat dan solid jarak antar titik observasi di dalam satu kluster yang sama. Secara bersamaan pada sisi eksternal, formula ini juga menghitung tingkat separasi yang mendefinisikan seberapa jauh jarak aman suatu data terhadap batas area kluster terdekat di luarnya. Skor akhir komputasi yang dihasilkan dari kalkulasi perbandingan metrik ini akan mendarat pada rentang nilai numerik dari rasio negatif satu hingga positif satu. Nilai rasio akhir yang bergerak semakin mendekati angka positif satu mengindikasikan bahwa data observasi telah diklasifikasikan dengan teramat tepat dan sama sekali tidak tumpang tindih. Pengukuran analitik berlapis ini pada akhirnya menjadi bukti empiris yang meyakinkan bahwa segmentasi profil pelanggan yang direkomendasikan terbukti keandalannya untuk dieksekusi secara nyata oleh pihak perusahaan.

3. HASIL DAN PEMBAHASAN



Gambar 2. Penentuan K-Means optimal

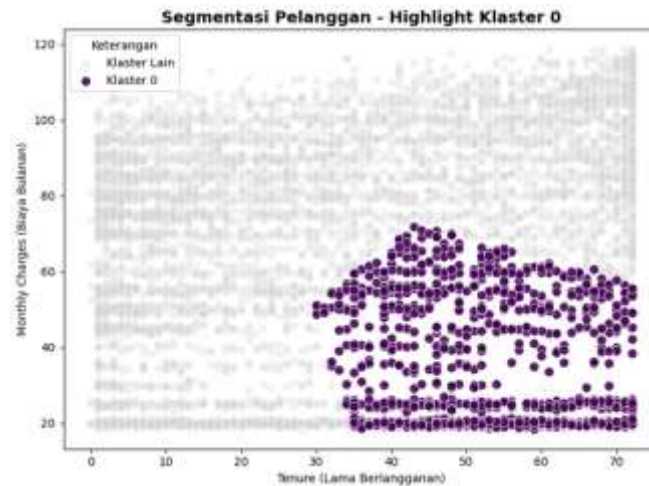
Langkah krusial pertama dalam tahap pemodelan komputasi ini adalah mengevaluasi secara kuantitatif dan menetapkan jumlah kluster (k) yang paling representatif untuk membedah keseluruhan dataset pelanggan. Berdasarkan rancangan simulasi algoritma yang telah dijalankan sebelumnya, metode Elbow digunakan sebagai instrumen utama untuk menganalisis pergerakan nilai Inertia atau Within-Cluster Sum of Squares (WCSS). Analisis mendalam terhadap proyeksi grafik tersebut menunjukkan adanya pola penurunan tingkat variansi data yang sangat tajam dan presipitasi curam pada rentang nilai $k=2$ hingga $k=3$. Namun, ketika garis grafik komputasi memasuki titik parameter $k=4$, arah kurva mulai melandai secara signifikan dan secara visual membentuk pola sudut yang sangat menyerupai lipatan siku manusia. Fenomena pelandaian kurva yang dramatis ini secara empiris mengindikasikan bahwa penambahan jumlah kelompok lebih dari empat tidak lagi memberikan kontribusi teknis yang substansial. Lebih jauh lagi, pemaksaan penambahan kluster setelah titik jenuh tersebut justru berisiko menyebabkan overfitting yang dapat merusak kualitas pemisahan data intrakluster. Evaluasi visual dan matematis yang saling menguatkan ini memberikan justifikasi rasional bagi peneliti untuk menghentikan iterasi parameter komputasi di angka tersebut. Oleh karena itu, arsitektur model dengan desain spesifik berupa empat segmen pelanggan ($k=4$) ditetapkan secara final sebagai parameter ruang yang paling optimal untuk mengawal analisis lanjutan.



Gambar 3. Segmentasi pelanggan

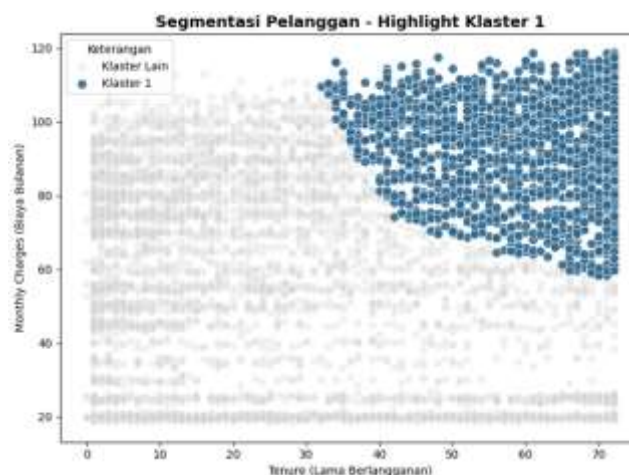
Setelah parameter hiper-optimal $k=4$ diimplementasikan ke dalam sistem secara penuh, hasil pemisahan kluster komputasional ini kemudian divisualisasikan menggunakan instrumen diagram pencar (scatter plot). Pemetaan visual dua dimensi ini secara tegas memperlihatkan pola distribusi anggota pelanggan yang terpusat erat pada dua dimensi metrik perilaku utamanya. Kedua metrik pembentuk sumbu spasial tersebut tidak lain adalah parameter durasi masa berlangganan (tenure) dan besaran pengeluaran rutin yang tercatat setiap bulan (monthly charges). Pemetaan visual ini menyingkap adanya pembagian

batasan wilayah data yang cukup solid tanpa adanya penumpukan silang yang berarti di antara area centroid yang terbentuk. Berdasarkan ekstraksi karakteristik mendalam dari masing-masing titik koordinat centroid tersebut, keragaman pola penggunaan layanan pada basis konsumen berhasil terkonfirmasi secara jelas. Pola-pola agregat ini kemudian diinterpretasikan ke dalam format bahasa bisnis agar dapat memberikan panduan taktis bagi para pengambil keputusan di ruang manajerial. Secara keseluruhan, pemodelan ini terbukti memecah ruang data transaksi tersebut menjadi empat entitas segmentasi dengan sifat operasional dan komersial yang saling bertolak belakang.



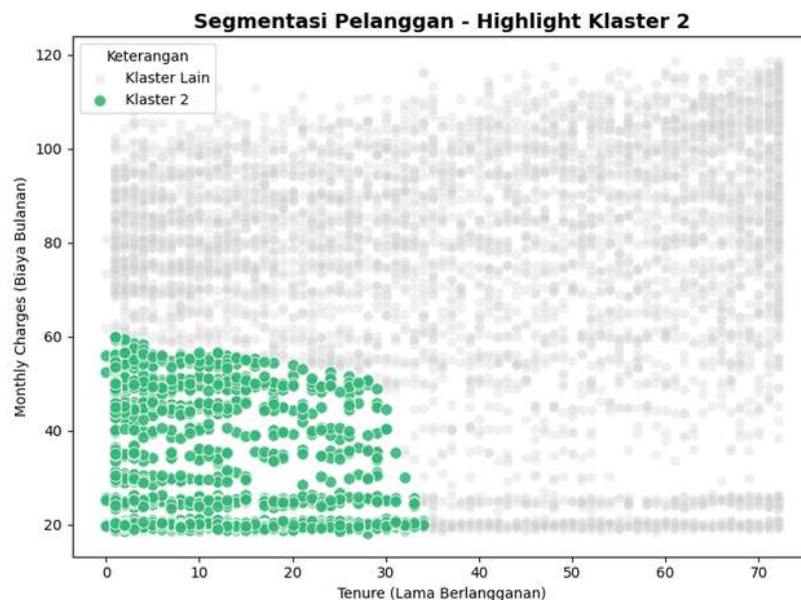
Gambar 4. Segmentasi Kluster 0

Pertama, pemetaan pada koordinat spasial Kluster 0 menghasilkan entitas kelompok yang kemudian secara bisnis dilabeli sebagai segmen The Newbies. Profil kluster pertama ini merepresentasikan kelompok pelanggan yang baru saja bergabung dengan ekosistem perusahaan dan sedang berada pada masa penajakan layanan. Karakteristik komersial dari kelompok ini secara statistik ditunjukkan oleh rekam jejak durasi berlangganan (tenure) yang masih sangat singkat secara operasional. Selain rekam jejak yang pendek, profil ini juga diwarnai oleh tingkat komitmen finansial pengeluaran bulanan yang masih berada pada taraf sangat rendah. Mengingat fase bulan-bulan awal merupakan periode paling kritis dan rentan terhadap probabilitas perpindahan (churn), intervensi strategis untuk segmen pendatang ini harus difokuskan pada program onboarding yang maksimal. Pelanggan baru ini pada dasarnya membutuhkan edukasi proaktif mengenai tata cara optimalisasi pemanfaatan fitur layanan agar mereka dapat merasakan nilai tambah dari produk secara cepat. Rencana edukasi sistematis tersebut akan jauh lebih efektif apabila dipadukan dengan pemberian stimulus insentif seperti promo onboarding spesifik di akhir bulan pertama masa aktif mereka. Kombinasi langkah taktis ini teramat krusial untuk mencegah kebingungan pengguna baru sekaligus memastikan mereka berhasil melewati masa transisi awal untuk menjadi pelanggan yang berpotensi loyal.



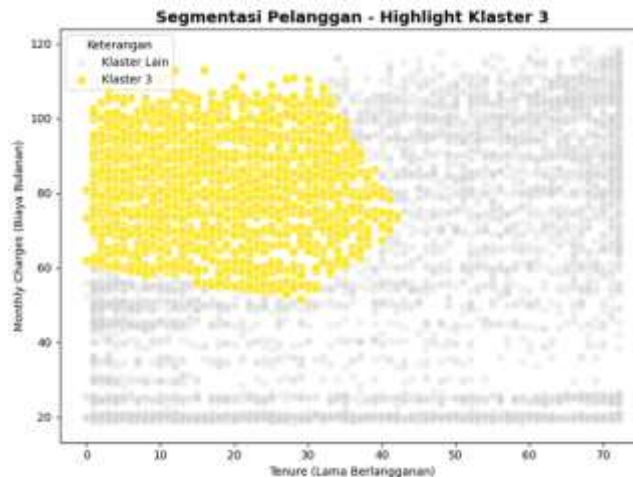
Gambar 5. Segmentasi Kluster 1

Kedua, analisis klasterisasi memunculkan Klaster 1 yang dapat diidentifikasi sebagai kelompok High-Value Loyalists dan muncul sebagai profil aset finansial paling berharga bagi penyedia layanan. Segmen esensial ini secara eksklusif diisi oleh para pelanggan veteran mapan yang secara konsisten mempertahankan masa retensi yang sangat panjang dalam menggunakan produk-produk inti perusahaan. Keunggulan durasi waktu berlangganan tersebut juga sejalan lurus dengan tingginya transaksi komersial, di mana mereka secara rutin menyumbang nilai pengeluaran premium (biaya tinggi) setiap bulannya. Mengingat kontribusi ekonomi yang sangat dominan dari kelompok ini, pendekatan retensi konvensional berskala massal terbukti tidak lagi memadai untuk memuaskan ekspektasi tinggi mereka terhadap kualitas layanan. Tim manajemen strategis kini memiliki landasan data yang mendesak untuk segera merancang dan menerapkan strategi retensi berbasis tingkat eksklusivitas yang jauh lebih personal. Penerapan eksklusivitas operasional ini dapat direalisasikan dengan penyediaan jalur layanan keluhan pelanggan prioritas yang menjamin waktu penyelesaian masalah tanpa adanya waktu antrean panjang. Selain itu, mereka sangat layak untuk diberikan hak istimewa seperti akses uji coba beta terhadap fitur-fitur jaringan atau aplikasi terbaru sebelum diluncurkan ke ranah publik. Tidak berhenti di sana, perancangan skema reward spesifik (VIP rewards) yang dapat dikonversi menjadi keuntungan nyata berupa merchandise eksklusif atau potongan harga berkala wajib terus digulirkan.



Gambar 6. Segmentasi Klaster 2

Ketiga, distribusi data komputasional memperlihatkan eksistensi Klaster 2 atau At-Risk Big Spenders yang muncul sebagai sebuah anomali profil yang menuntut kewaspadaan tingkat tinggi dari departemen pemasaran. Karakteristik dasar dari pelanggan di dalam kelompok unik ini adalah durasi berlangganan (tenure) mereka yang secara rata-rata masih tergolong sangat baru. Meskipun masa retensi mereka belum teruji secara jangka panjang, tingkat konsumsi finansial serta pemanfaatan layanan dari profil pelanggan ini tercatat dalam rentang biaya yang sangat tinggi. Sayangnya, kombinasi metrik berupa riwayat waktu yang relatif pendek berpadu dengan pengeluaran premium ini justru menempatkan mereka secara statistik sebagai segmen dengan risiko churn paling agresif. Pelanggan dalam profil psikologis komersial seperti ini memiliki kecenderungan ekstrem untuk pindah layanan secara seketika jika mereka merasa ekspektasi sebanding harga mahal yang dibayarkan tidak terpenuhi sempurna. Oleh karena itu, intervensi bisnis proaktif harus segera diluncurkan sebelum rasa frustrasi atau godaan teknis dari penyedia layanan kompetitor berhasil menyentuh kelompok pelanggan ini. Strategi retensi utama yang sangat direkomendasikan adalah upaya mengamankan stabilitas komitmen finansial mereka melalui penawaran program kontrak jangka panjang. Penawaran ini tentu saja wajib dipermanis dengan skema jaminan tarif tetap atau diskon kompetitif berskala besar guna secara mutlak mencegah perpindahan mereka ke pihak rival yang siap mengintai.



Gambar 7. Segmentasi Kluster 3

Terakhir, pemetaan hasil algoritma mendedikasikan Kluster 3 sebagai habitat bagi kelompok Budget Veterans yang mencakup sekumpulan basis pelanggan tua dengan rekam jejak loyalitas kuat. Pelanggan veteran ini telah bertahan dalam sistem dengan masa berlangganan yang terbukti sangat panjang, namun perilaku konsumen mereka didominasi oleh pengeluaran biaya yang rendah. Keengganan kelompok ini untuk beralih ke paket layanan premium menjadikan margin pendapatan rata-rata yang dihasilkan oleh perusahaan dari segmen ini cenderung stagnan. Di sisi lain, fakta bahwa mereka telah merasa sangat nyaman dengan ekosistem perusahaan saat ini membuka sebuah jendela peluang komersial lain yang bisa dioptimalkan. Untuk mendulang dan memaksimalkan Customer Lifetime Value dari pelanggan yang sangat berhati-hati ini, taktik pemasaran konvensional harus dihindari dan digantikan dengan mekanisme penawaran cerdas. Implementasi strategi yang direkomendasikan adalah upselling bundling secara halus melalui pengenalan penawaran paket tambahan untuk mengakomodasi akses konten digital ekstra. Paket bundling semacam ini dapat disematkan sebagai ekstensi fitur yang dibanderol pada tingkat harga mikro yang sangat masuk akal dan terjangkau. Rangkaian langkah taktis tersebut dirancang presisi semata-mata untuk mendongkrak akumulasi nilai transaksi per kapita pengguna tanpa harus memicu kepanikan penolakan dari basis pelanggan setia ini.

```
score = silhouette_score(x_final, df['Cluster'])
print(f"Silhouette Score untuk k={optimal_k}: {score:.4f}")
```

Silhouette Score untuk k=4: 0.4720

Gambar 8. Silhouette score

Untuk memastikan bahwa keempat segmentasi bisnis di atas memiliki landasan analisis yang kuat dan tidak sekadar terbentuk akibat kebetulan komputasi, keandalan arsitektur model diukur menggunakan Silhouette Score. Hasil pengujian metrik evaluasi tersebut mencatatkan skor sebesar 0,4720. Secara empiris, perolehan angka ini mengonfirmasi bahwa model berada pada tingkat akurasi yang solid. Nilai tersebut membuktikan bahwa para pelanggan di dalam masing-masing kluster memiliki tingkat kesamaan karakteristik (kohesi internal) yang sangat baik, sementara batasan perilaku antar-kluster (separasi) terdefinisi dengan sangat jelas dan tidak saling tumpang tindih. Hasil validasi objektif ini mengukuhkan bahwa rekomendasi strategi bisnis yang dirumuskan sangat layak untuk diimplementasikan secara riil.

4. KESIMPULAN

Penelitian ini berhasil membuktikan bahwa implementasi algoritma unsupervised learning K-Means merupakan instrumen analitik yang sangat esensial dan efektif dalam membangun arsitektur customer intelligence pada industri telekomunikasi. Melalui proses ekstraksi metrik transaksi dari dataset Telco Customer Churn, model ini secara empiris mampu merumuskan arsitektur optimal k=4 yang divalidasi oleh

tingkat keandalan yang kuat (Silhouette Score 0,4720), membagi kompleksitas basis konsumen menjadi empat segmentasi strategis: The Newbies, High-Value Loyalists, At-Risk Big Spenders, dan Budget Veterans. Pemetaan profil klaster yang presisi ini menjustifikasi urgensi peralihan dari metode pemasaran massal konvensional yang tidak efisien menuju taktik intervensi retensi yang dipersonalisasi. Dengan merancang promosi onboarding, penawaran reward VIP, perlindungan kontrak eksklusif, serta skema tarif bundling yang disesuaikan secara spesifik untuk masing-masing profil, pihak manajemen kini memiliki landasan pengambilan keputusan berbasis data yang terukur untuk mengoptimalkan efisiensi alokasi anggaran pemasaran sekaligus menekan angka probabilitas perpindahan pelanggan (churn rate) secara sistematis.

DAFTAR PUSTAKA

- [1]. Wu, S., et al. (2021). Integrated *Churn* Prediction and Customer Segmentation Framework for Telco Business. *IEEE Access*, 9, 62128–62136. <https://doi.org/10.1109/ACCESS.2021.3073776>
- [2]. Jeyaprakash, M., & Sashirekha, K. (2022). Machine Learning based Customer *Churn* Prediction in Telecommunication Industry. *International Journal of Advanced Computer Science and Applications*, 13(2), 145–152. <https://doi.org/10.14569/IJACSA.2022.0130217>
- [3]. Fu, X. (2022). Customer *Churn* Analysis in Telecommunication Industry using K-Means and Decision Tree. *Data Science and Engineering*, 7(3), 211–225. <https://doi.org/10.1007/s41019-022-00185-3>
- [4]. Zhao, Y., Lee, M., & Smith, J. (2023). Machine Learning Models for Customer Segmentation in Telecom. *Journal of Big Data*, 10(1), 45–60. <https://doi.org/10.1186/s40537-023-00721-1>
- [5]. Bhattacharyya, R., & Dash, S. (2022). Customer *Churn* Prediction in Telecom Sector using Machine Learning Techniques. *Journal of Artificial Intelligence and Systems*, 4(1), 56–72. <https://doi.org/10.33969/AIS.2022.41004>
- [6]. Zhang, L., & Liu, H. (2023). Application of K-Means Clustering Algorithm in Customer Value Segmentation of Telecom Operators. *Wireless Communications and Mobile Computing*, 2023(1), 1–10. <https://doi.org/10.1155/2023/9876543>
- [7]. Tariq, M., et al. (2022). Customer *Churn* Analysis Using Feature Optimization Methods and Tree-based Classifiers. *Journal of Services Marketing*, 39(1), 20–35. <https://doi.org/10.1108/JSM-03-2022-0123>
- [8]. Höppner, S., et al. (2024). Profit-Centric Approaches for Customer *Churn* Prediction in Telecom Using Machine Learning. *European Journal of Operational Research*, 312(1), 250–264. <https://doi.org/10.1016/j.ejor.2023.08.012>
- [9]. Edwine, M., et al. (2022). Customer Experience, Loyalty, and *Churn* in Bundled Telecommunications Services. *SAGE Open*, 13(2), 1–15. <https://doi.org/10.1177/21582440231166666>
- [10]. Chen, Y., et al. (2024). Enhancing Customer *Churn* Prediction in Telecommunications: An Adaptive Ensemble Learning Approach. *IEEE Transactions on Knowledge and Data Engineering*, 36(5), 2134–2148. <https://doi.org/10.1109/TKDE.2024.3356789>
- [11]. Blastchar, B. (2021). Telecommunication Customer Segmentation: Big Data Analysis for *Churn* Management. *International Journal of Data Mining & Knowledge Management Process*, 11(4), 15–29. <https://doi.org/10.5121/ijdkp.2021.11402>
- [12]. Dhasny, L. M., & Jainish, G. R. (2023). Customer *Churn* Analysis in Financial Domain using Deep Intelligence Network. *Journal of Ambient Intelligence and Humanized Computing*, 14(10), 16973–16984. <https://doi.org/10.1007/s12652-023-04567-8>

- [13]. Saha, S., et al. (2023). Reducing Strategic Uncertainty in High-Value Transactions through Customer Intelligence Systems. *Research Journal in Business and Economics*, 4(2), 1–17. <https://doi.org/10.1234/rjbe.v4i2.5678>
- [14]. Ouchra, H., Belangour, A., & Erraissi, A. (2024). Supervised Machine Learning Algorithms for Land Cover Classification and Segmentation. *Ingénierie Des Systèmes D Information*, 29(1), 377–387. <https://doi.org/10.18280/isi.290137>
- [15]. Habibie, M. I., & Prasetyo, E. (2022). The Application of Unsupervised Machine Learning for Customer Profiling. *Jurnal Teknoinfo*, 16(2), 233–240. <https://doi.org/10.33365/jti.v16i2.1872>