

Penerapan Algoritma K-Means Berbasis Model RFM untuk Klasterisasi Anggota di Perpustakaan Daerah Kuala Kapuas

Winda Astria Nuansa Saputri^{1*}, Windarsyah², Ayu Ahadi Ningrum³

^{1,2,3}S1 Informatika, Universitas Muhammadiyah Banjarmasin, Barito Kuala, Indonesia

Email: ^{1*}winda_astria_2255201110020@umbjm.ac.id ²windarsyah@umbjm.ac.id, ³ayuahadi@umbjm.ac.id

Email Penulis Korespondensi: ^{1*}winda_astria_2255201110020@umbjm.ac.id

Info Artikel	Abstrak
Kata Kunci: Data Mining K-Means Perpustakaan Daerah RFM Segmentasi Anggota	Perpustakaan daerah belum memiliki mekanisme berbasis data untuk membedakan anggota berdasarkan pola kunjungan dan peminjaman, sehingga layanan masih diberikan secara seragam kepada seluruh anggota. Penelitian ini bertujuan menerapkan model <i>Recency, Frequency, Monetary</i> (RFM) dikombinasikan dengan Algoritma K-Means untuk menghasilkan segmentasi berbasis data terhadap 1.430 anggota Dinas Kearsipan dan Perpustakaan Daerah Kuala Kapuas. Metode yang digunakan meliputi konstruksi variabel RFM dari data kunjungan dan peminjaman, normalisasi MinMaxScaler, penentuan jumlah kluster optimal menggunakan <i>Elbow Method</i> , implementasi K-Means, serta validasi kualitas kluster menggunakan <i>Silhouette Score</i> , <i>Davies-Bouldin Index</i> , dan <i>Calinski-Harabasz Index</i> . Hasil penelitian menunjukkan K=3 sebagai jumlah kluster optimal dengan nilai <i>Silhouette Score</i> sebesar 0,3478. Tiga kluster yang terbentuk adalah Anggota Aktif (199 anggota, 13,9%), Anggota Sedang (570 anggota, 39,9%), dan Anggota Pasif (661 anggota, 46,2%), dengan dimensi <i>Frequency</i> sebagai pembeda utama antar kluster. Anggota Pasif tercatat memiliki tingkat kepuasan layanan yang tinggi meskipun jarang berkunjung, yang mengindikasikan hambatan akses sebagai penyebab kepasifan. Berdasarkan profil tiap kluster, dirumuskan tiga strategi layanan terdiferensiasi berupa retensi untuk Anggota Aktif, peningkatan keterlibatan untuk Anggota Sedang, dan reaktivasi untuk Anggota Pasif.
	Abstract
Keywords: Data Mining K-Means Member Segmentation RFM Regional Library	Regional libraries lack data-driven mechanisms to differentiate members based on visit and borrowing patterns, resulting in uniform services delivered to all members regardless of their characteristics. This study aims to apply the <i>Recency, Frequency, Monetary</i> (RFM) model combined with the K-Means Algorithm to generate data-driven segmentation of 1,430 registered members at the Kuala Kapuas Regional Library and Archives Office. The method involves RFM variable construction from visit and borrowing data, MinMaxScaler normalization, optimal <i>cluster</i> determination using the <i>Elbow Method</i> , K-Means implementation, and cluster quality validation using <i>Silhouette Score</i> , <i>Davies-Bouldin Index</i> , and <i>Calinski-Harabasz Index</i> . Results indicate K=3 as the optimal number of clusters with a <i>Silhouette Score</i> of 0.3478. Three <i>clusters</i> were identified: Active Members (199 members, 13.9%), Moderate Members (570 members, 39.9%), and Passive Members (661 members, 46.2%), with the <i>Frequency</i> dimension as the primary differentiator among <i>clusters</i> . Passive Members reported high service satisfaction despite infrequent visits, indicating that access barriers rather than dissatisfaction drive passivity. Based on each <i>cluster</i> profile, three differentiated service strategies were formulated: retention for Active Members, engagement improvement for Moderate Members, and reactivation for Passive Members.

JuKSIT is licensed under a Creative Commons Attribution-Share Alike 4.0 International License



1. PENDAHULUAN

Perpustakaan daerah memegang peran strategis sebagai pusat literasi masyarakat dan penyedia akses informasi publik. Dinas Kearsipan dan Perpustakaan Daerah Kuala Kapuas saat ini mengelola 1.430 anggota terdaftar dengan karakteristik dan pola pemanfaatan layanan yang beragam. Namun demikian, program layanan yang ditawarkan masih dirancang secara seragam tanpa mempertimbangkan perbedaan perilaku antar kelompok anggota, sehingga intervensi yang diberikan tidak selalu tepat sasaran. Kondisi ini mengakibatkan sebagian anggota, terutama yang tergolong pasif,

tidak terlayani secara optimal meskipun data keanggotaan yang memuat riwayat kunjungan, peminjaman, dan kategori koleksi telah tersedia [1]. Tantangan serupa dihadapi oleh mayoritas perpustakaan daerah di Indonesia, di mana keterbatasan sumber daya dan anggaran operasional mendorong pengelola mengadopsi pendekatan layanan yang generik dan tidak terpersonalisasi [2]. Akibatnya, potensi data operasional yang telah terakumulasi selama bertahun-tahun belum dimanfaatkan sebagai basis pengambilan keputusan berbasis bukti (*evidence-based decision making*) [3]. Padahal, data kunjungan dan peminjaman anggota menyimpan pola perilaku yang kaya dan dapat dieksploitasi untuk merancang intervensi layanan yang lebih presisi dan berdampak [4].

Dalam konteks tersebut, pendekatan pengelompokan berbasis kemiripan karakteristik melalui teknik *clustering* dalam *data mining* mampu mengidentifikasi pola perilaku pengguna secara otomatis tanpa membutuhkan perubahan pada infrastruktur sistem yang telah tersedia [5]. Di antara berbagai pendekatan segmentasi berbasis perilaku, model RFM (*Recency, Frequency, Monetary*) menonjol karena kemampuannya mengubah data transaksi menjadi representasi numerik yang mencerminkan tiga aspek perilaku pengguna, yakni jarak waktu sejak interaksi terakhir, tingkat keseringan interaksi, serta besaran volume aktivitas yang terekam dalam sistem [6]. Keunggulan model RFM terletak pada kemampuannya menghasilkan gambaran perilaku yang mudah ditafsirkan dan dapat segera diterapkan hanya dengan memanfaatkan data transaksi yang sudah ada, tanpa bergantung pada kelengkapan data demografis pengguna [7]. Keunggulan ini menjadikan RFM sangat relevan dalam konteks institusi publik seperti perpustakaan daerah, di mana ketersediaan data demografis anggota seringkali tidak lengkap atau tidak terstruktur, sementara data transaksi operasional seperti tanggal kunjungan, jumlah buku dipinjam, dan kategori koleksi yang diakses tersedia secara konsisten dalam sistem pencatatan yang ada [8].

Beberapa penelitian terdahulu telah menerapkan kombinasi RFM dan K-Means dalam berbagai domain. Studi empiris yang dilakukan oleh Akande dkk. mengonfirmasi bahwa kombinasi model RFM dengan algoritma K-Means berhasil menghasilkan pengelompokan pelanggan yang lebih tajam sehingga strategi intervensi yang dirancang menjadi jauh lebih tepat sasaran [9]. Selanjutnya, Aqil, Anam, dan Sianipar menunjukkan bahwa pendekatan serupa pada data transaksi ritel menghasilkan segmen yang kohesif dan memiliki implikasi praktis langsung bagi pengambil keputusan bisnis [10]. Dalam konteks perpustakaan, Putri, Nurhasanah, dan Perangin-Angin menerapkan metode K-Means *clustering* dan aturan asosiasi untuk menghasilkan rekomendasi bahan bacaan yang lebih sesuai dengan preferensi pengunjung perpustakaan [11]. Sementara itu, Zabidi menerapkan algoritma K-Means untuk mengelompokkan koleksi perpustakaan sehingga dapat membantu pengelolaan dan pemanfaatan koleksi secara lebih efektif [12]. Meskipun demikian, penerapan model RFM secara khusus dalam konteks perpustakaan daerah di Indonesia sebagai kerangka kerja segmentasi anggota yang dapat langsung dioperasionalkan masih sangat terbatas dalam literatur yang ada [13]. Sebagian besar studi yang ada berfokus pada perpustakaan perguruan tinggi atau konteks komersial, sehingga terdapat celah penelitian (*research gap*) yang signifikan terkait validitas dan efektivitas pendekatan segmentasi berbasis RFM pada perpustakaan daerah dengan karakteristik anggota yang lebih heterogen dan motivasi kunjungan yang lebih bervariasi [14]. Penelitian ini hadir untuk mengisi celah tersebut dengan mengadaptasi kerangka kerja RFM secara kontekstual pada setting perpustakaan daerah, sekaligus menawarkan pendekatan yang tidak memerlukan investasi sistem baru sehingga dapat langsung diimplementasikan dalam kondisi keterbatasan infrastruktur yang lazim dijumpai pada institusi pemerintah daerah.

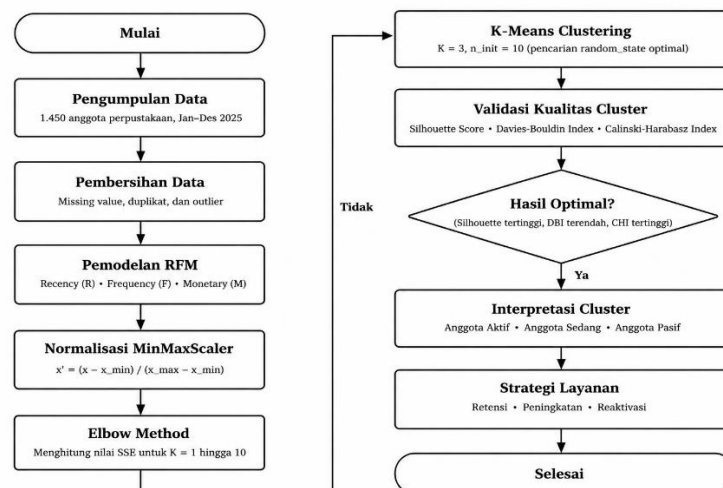
Pemilihan algoritma K-Means dalam penelitian ini didasarkan pada karakteristiknya yang terbukti handal dalam memproses kumpulan data berskala besar dengan biaya komputasi yang relatif ringan dibandingkan algoritma *clustering* lainnya [15]. Algoritma seperti DBSCAN sangat bergantung pada parameter epsilon yang sulit ditentukan secara objektif, sementara K-Medoids memiliki beban komputasi $O(n^2)$ yang kurang praktis untuk data besar. K-Means hadir sebagai solusi tengah yang menyeimbangkan efisiensi proses dengan kemudahan interpretasi hasil, khususnya pada dataset keanggotaan perpustakaan yang berjumlah ribuan rekaman [15]. Untuk memastikan objektivitas ilmiah, jumlah *cluster* optimal ditentukan menggunakan *Elbow Method* yang menghitung nilai *Sum of Squared Errors* (SSE) pada rentang $K=1$ hingga $K=10$ [16], dan kualitas hasil klusterisasi divalidasi menggunakan *Silhouette Score* metrik evaluasi yang mengukur kohesi internal sekaligus separasi antar *cluster* [17]. Kombinasi ketiga komponen ini – model RFM, Algoritma K-Means, dan validasi *Silhouette Score* – membentuk kerangka metodologis yang komprehensif, dapat direproduksi, dan dapat dipertanggungjawabkan secara akademis.

Berdasarkan latar belakang yang telah diuraikan, penelitian ini bertujuan untuk menerapkan pemodelan RFM pada data keanggotaan perpustakaan guna menghasilkan representasi kuantitatif perilaku 1.430 anggota terdaftar,

mengimplementasikan Algoritma K-Means dengan *Elbow Method* untuk menentukan jumlah *cluster* optimal dan mengklasterisasi seluruh anggota ke dalam segmen yang bermakna, memvalidasi kualitas hasil klasterisasi menggunakan *Silhouette Score*, *Davies-Bouldin Index*, dan *Calinski-Harabasz Index* sebagai bukti empiris yang dapat dipertanggungjawabkan secara ilmiah, dan merumuskan rekomendasi strategi layanan yang terdiferensiasi dan operasional untuk setiap segmen anggota, meliputi strategi retensi bagi anggota aktif, peningkatan keterlibatan bagi anggota sedang, dan reaktivasi bagi anggota pasif, yang dapat langsung diimplementasikan tanpa memerlukan investasi sistem baru.

2. METODOLOGI PENELITIAN

Studi ini dirancang menggunakan kerangka penelitian kuantitatif yang bertumpu pada teknik penambangan data, khususnya metode pengelompokan otomatis, yang diterapkan pada basis data keanggotaan Dinas Kearsipan dan Perpustakaan Daerah Kuala Kapuas. Seluruh komputasi diimplementasikan menggunakan Python dengan pustaka *scikit-learn*, *pandas*, *NumPy*, dan *matplotlib*. Tahapan penelitian disusun secara sistematis mulai dari proses pengumpulan data hingga perumusan strategi layanan. Alur lengkap penelitian ditunjukkan pada Gambar 1.



Gambar 1. Diagram Alir Tahapan Penelitian

2.1 Sumber Data

Data yang menjadi bahan analisis bersumber dari sistem informasi perpustakaan, mencakup 1.430 catatan anggota yang tercatat aktif selama periode Januari hingga Desember 2025. Atribut yang digunakan disajikan pada Tabel 1.

Tabel 1. Atribut Dataset dan Kegunaannya

Atribut	Keterangan	Digunakan untuk
Tanggal Kunjungan	Tanggal kunjungan terakhir anggota	Menghitung R
Jumlah Buku Dipinjam	Total buku dipinjam per kunjungan	Menghitung F
Jumlah Pinjam Kategori	Volume peminjaman per kategori koleksi	Menghitung M

Tahap pra-pemrosesan data mencakup tiga langkah utama yaitu penanganan nilai yang hilang, penghapusan entri ganda, serta penyeragaman format penulisan tanggal agar seluruh rekaman dapat diproses secara konsisten. Deteksi *outlier* menggunakan IQR [18] menunjukkan tidak ada nilai ekstrem yang perlu dieliminasi, sehingga seluruh 1.430 rekaman digunakan dalam analisis.

2.2 Pemodelan RFM

Model RFM mengkuantifikasi perilaku anggota melalui tiga dimensi [19]: *Recency* (R), yaitu selang hari antara tanggal kunjungan terakhir dengan tanggal referensi 26 Desember 2025, *Frequency* (F), yaitu total buku yang dipinjam; dan *Monetary* (M), yaitu total volume peminjaman lintas kategori koleksi sebagai proksi tingkat keterlibatan *engagement* [20]. Dimensi M diadaptasi dari nilai finansial karena perpustakaan tidak menghasilkan transaksi moneter, sehingga intensitas aktivitas peminjaman digunakan sebagai ukuran kuantitatif keterlibatan anggota [21].

Pendekatan ini telah terbukti efektif pada layanan non-komersial, di mana volume aktivitas pengguna dapat merepresentasikan tingkat *engagement* secara objektif [22]. Selanjutnya, ketiga dimensi RFM digunakan sebagai fitur masukan bagi algoritma K-Means untuk mengelompokkan anggota yang memiliki karakteristik perilaku serupa ke dalam klaster yang sama [23]. Dengan representasi tersebut, segmentasi yang dihasilkan menjadi lebih mudah diinterpretasikan dan dapat dimanfaatkan sebagai dasar penyusunan strategi layanan perpustakaan tanpa memerlukan atribut demografis tambahan [24].

2.3 Normalisasi dan Penentuan K Optimal

Untuk memastikan tidak ada satu variabel pun yang mendominasi proses penghitungan jarak antar titik data, ketiga dimensi RFM diseragamkan rentang nilainya menggunakan transformasi *MinMaxScaler* [25], menggunakan persamaan (1).

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

di mana x adalah nilai asal, x_{min} dan x_{max} adalah nilai minimum dan maksimum aktual masing-masing dimensi. Batas aktual: R(0–1.088), F(1–25), M(5–34). Penentuan jumlah klaster yang paling sesuai dilakukan dengan mengevaluasi nilai SSE pada rentang K=1 sampai K=10 melalui pendekatan *Elbow Method*, di mana titik perlambatan penurunan SSE dijadikan acuan penetapan K optimal (2) [26]:

$$SSE = \sum_{i=1}^K \sum_{x \in C_i} \|x - c_i\|^2 \quad (2)$$

di mana C_i adalah himpunan anggota *cluster* ke- i dan c_i adalah centroidnya. Nilai K pada titik "siku" di mana penurunan SSE mulai melandai ditetapkan sebagai K optimal.

2.4 Implementasi K-Means dan Validasi

K-Means mengklasterisasi data *Euclidean Distance* di hitung menggunakan persamaan (3) [27]:

$$d(x, C_j) = \sqrt{\sum_{i=1}^n (x_i - c_{ij})^2} \quad (3)$$

dimana x_i adalah nilai fitur ke- i dari titik data dan c_{ij} adalah nilai centroid *cluster* ke- j pada fitur ke- i . Implementasi menggunakan parameter K=3, *random_state*=0, *n_init*=10.

Penamaan cluster dilakukan berdasarkan nilai Total Skor yang diperoleh dari kombinasi nilai *Recency*, *Frequency*, dan *Monetary* yang telah dinormalisasi menggunakan Persamaan (4):

$$\text{Total Skor} = (1 - R') + F' + M' \quad (4)$$

Validasi kualitas *cluster* menggunakan tiga metrik komplementer: *Silhouette Score* dihitung menggunakan persamaan (5) [28]:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (5)$$

dimana $a(i)$ adalah rata-rata jarak titik data terhadap anggota dalam *cluster* yang sama, dan $b(i)$ adalah rata-rata jarak terhadap *cluster* terdekat yang berbeda; *Davies-Bouldin Index* mengukur pemisahan antar *cluster* dengan nilai lebih

kecil menandakan kualitas lebih baik; serta *Calinski-Harabasz Index* mengevaluasi kekompakan dan separasi *cluster* dengan nilai lebih besar menandakan hasil yang lebih baik [29].

2.5 Perumusan Strategi Layanan

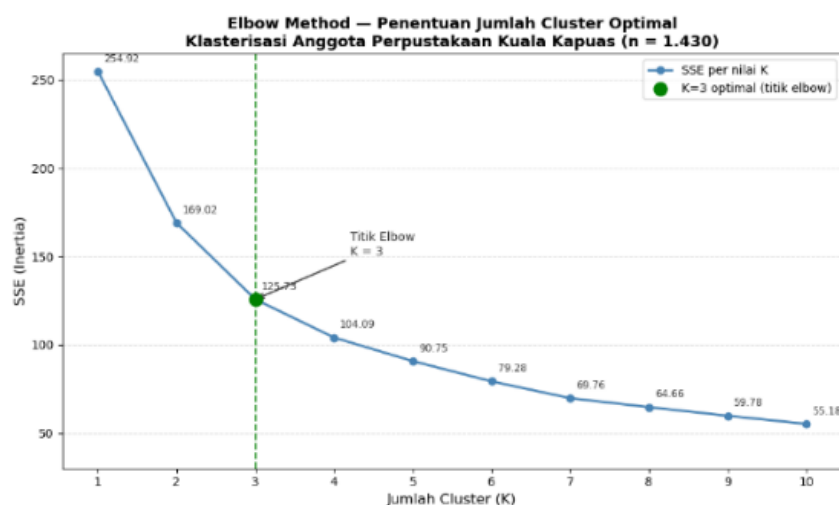
Perumusan rekomendasi layanan dilakukan dengan mengacu pada profil masing-masing klaster yang terbentuk, sehingga setiap saran yang dihasilkan memiliki sifat praktis dan dapat dijalankan langsung oleh pihak pengelola perpustakaan tanpa memerlukan pengembangan atau modifikasi sistem apapun, mencakup retensi untuk Anggota Aktif, peningkatan keterlibatan untuk Anggota Sedang, dan reaktivasi untuk Anggota Pasif [30]. Pendekatan ini sejalan dengan penelitian yang menegaskan bahwa hasil *clustering* dapat dijadikan dasar penyusunan strategi manajemen layanan yang lebih tepat sasaran sesuai karakteristik tiap segmen [31]. Dalam konteks perpustakaan, segmentasi berbasis data keanggotaan terbukti mampu menghasilkan rekomendasi intervensi yang spesifik dan aplikatif untuk setiap kelompok pengguna [32]. Selain itu, pendekatan klasterisasi pada data keanggotaan perpustakaan dapat dimanfaatkan secara langsung untuk menentukan prioritas dan lokasi strategi layanan yang lebih efektif [33].

3. HASIL DAN PEMBAHASAN

3.1 Penentuan Jumlah Cluster Optimal

3.1.1 Hasil *Elbow Method*

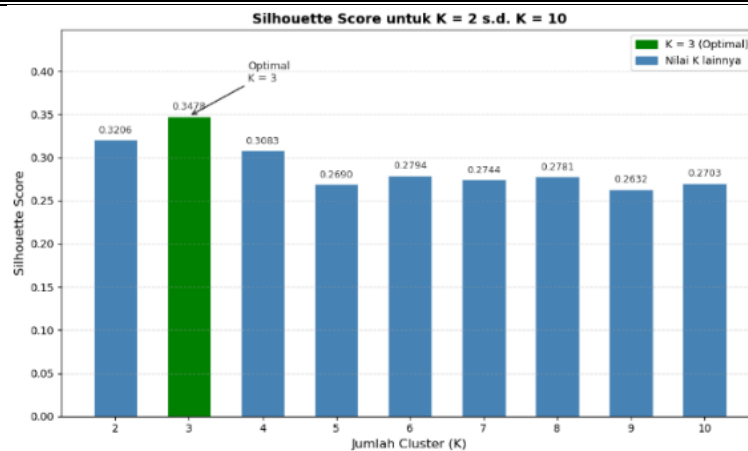
Penentuan jumlah *cluster* optimal dilakukan dengan menghitung nilai SSE (persamaan 2) pada rentang $K=2$ hingga $K=10$ menggunakan data R' , F' , M' yang telah dinormalisasi persamaan (1). Hasil perhitungan disajikan pada Gambar 2.



Gambar 2. Grafik *Elbow Method*

Berdasarkan Gambar 2, penurunan SSE paling signifikan terjadi dari $K=2$ (169,02) ke $K=3$ (125,73), yaitu sebesar 43,29 poin. Setelah $K=3$, laju penurunan SSE melandai secara konsisten dari $K=3$ ke $K=4$ hanya turun 21,64 poin, dan terus berkurang hingga hanya 4,60 poin dari $K=9$ ke $K=10$. Titik "siku" (*elbow*) secara visual berada pada $K=3$, mengindikasikan bahwa penambahan *cluster* di atas $K=3$ tidak memberikan peningkatan kualitas yang sebanding dengan peningkatan kompleksitas model.

3.1.2 Validasi Kualitas Cluster

**Gambar 3.** Grafik *Silhouette Score*

Validasi kualitas *cluster* menggunakan tiga metrik yang saling melengkapi sebagaimana ditunjukkan pada Gambar 3 dan Tabel 2. *Silhouette Score* (persamaan 5) sebesar 0,3478 merupakan nilai tertinggi di antara seluruh alternatif K, menunjukkan bahwa sebagian besar anggota lebih mirip dengan *cluster* asalnya dibandingkan *cluster* lain. Meski belum tergolong sangat tinggi, nilai ini tetap dapat diterima mengingat data perilaku manusia secara alami memiliki karakteristik yang saling tumpang tindih, sejalan dengan Ardiansyah et al. yang menegaskan bahwa nilai *Silhouette Score* di kisaran 0,3–0,5 tetap menghasilkan segmen bermakna selama didukung metrik validasi komplementer [28]. Hal ini dikonfirmasi oleh *Davies-Bouldin Index* (1,0381) sebagai nilai terendah di antara seluruh K yang diuji, menandakan jarak antar-centroid relatif lebih besar dibandingkan dispersi di dalam *cluster* sehingga pemisahan antarkelompok cukup baik, serta *Calinski-Harabasz Index* (733,07) sebagai nilai tertinggi di antara seluruh alternatif jumlah *cluster*, sehingga keputusan penetapan K=3 dapat dipertanggung jawabkan secara ilmiah [34].

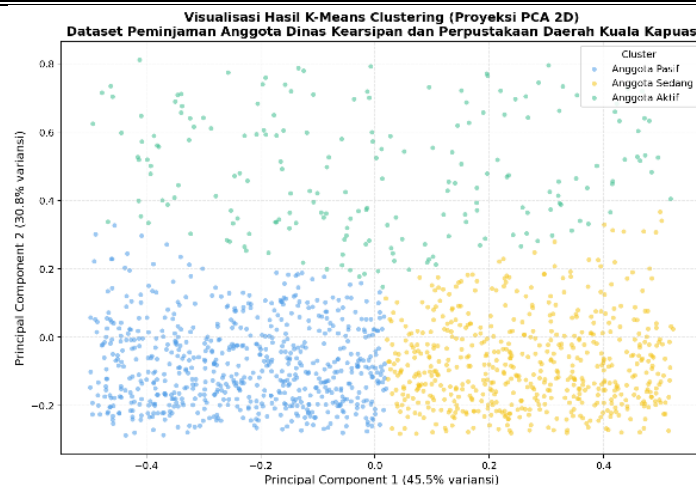
3.2 Hasil K-Means Clustering

3.2.1 Centroid Akhir dan Penamaan *Cluster*

Setelah algoritma K-Means mengklusterisasi data berdasarkan jarak *Euclidean* (persamaan 3) dan konvergen dengan parameter K=3, *random_state*=0, dan *n_init*=10, diperoleh centroid akhir yang disajikan pada Tabel 3. Penamaan *cluster* dilakukan berdasarkan total skor $(1-R') + F' + M'$, yang mempertimbangkan arah R' secara benar nilai R' besar menandakan kunjungan lama, bukan aktif.

Tabel 2. Centroid Akhir K-Means dan Penamaan *Cluster*

Label <i>Cluster</i>	n	Centroid R'	Centroid F'	Centroid M'	Total Skor	Rank
Anggota Aktif	199	0,5104	0,7305	0,3322	1,5523	1
Anggota Sedang	570	0,2460	0,1740	0,2842	1,2122	2
Anggota Pasif	661	0,7431	0,1730	0,3082	0,7381	3

**Gambar 4.** Visualisasi PCA K-Means

Proyeksi PCA pada Gambar 4 menunjukkan tiga kelompok yang secara visual cukup terpisah meskipun terdapat irisan di beberapa area, yang wajar terjadi pada data perilaku pengguna dan konsisten dengan Silhouette Score sebesar 0,3478 yang mengindikasikan irisan marginal antarcluster. *Cluster* Anggota Aktif memiliki sebaran paling luas mencerminkan keberagaman profil pada dimensi R dan M meskipun sama-sama berintensitas peminjaman tinggi, sedangkan Anggota Sedang dan Pasif membentuk kelompok yang lebih rapat karena disatukan oleh karakteristik intensitas peminjaman yang rendah.

3.2.2 Distribusi dan Profil *Cluster*

Distribusi anggota ke tiga *cluster* dan profil variabel pendukungnya disajikan pada Tabel 3. Penamaan *cluster* dilakukan berdasarkan total skor (persamaan 4).

Tabel 3. Distribusi dan Profil *Cluster*

Label	n	%	Rata-rata R (hari)	Rata-rata F (buku)	Rata-rata M	Durasi (menit)
Anggota Aktif	199	13,9%	555,33	18,53	14,63	91,21
Anggota Sedang	570	39,9%	267,66	5,18	13,24	89,79
Anggota Pasif	661	46,2%	808,54	5,15	13,94	91,04
Total	1.430	100%	—	—	—	—

Temuan bahwa hampir separuh anggota (46,2%) tergolong pasif mengkonfirmasi urgensi penelitian ini layanan yang selama ini diberikan secara seragam terbukti tidak mampu mengaktifkan kelompok mayoritas tersebut. Distribusi ini juga memperlihatkan bahwa Anggota Aktif hanya berjumlah 13,9%, namun kelompok kecil inilah yang sesungguhnya menjadi tulang punggung pemanfaatan koleksi perpustakaan secara intensif.

3.3 Karakteristik Tiap *Cluster*

3.3.1 Anggota Aktif (n=199, 13,9%)

Anggota Aktif dicirikan oleh nilai F' tertinggi (0,7305) dengan rata-rata peminjaman 18,53 buku per kunjungan, hampir empat kali lipat klaster lainnya. Kategori koleksi favorit meliputi Pendidikan & Akademik, Agama & Spiritual, dan Kesehatan & Medis, dengan 58,3% berstatus Anggota Tetap yang mengindikasikan korelasi antara keanggotaan formal dan intensitas peminjaman. Tingkat kepuasan mencapai 74,4% (Puas dan Sangat Puas), sehingga strategi retensi lebih efisien secara biaya dibandingkan upaya reaktivasi.

3.3.2 Anggota Sedang (n=570, 39,9%)

Anggota Sedang memiliki R' terendah (0,2460) dengan rata-rata kunjungan terakhir 267,66 hari lalu, menandakan hubungan yang masih aktif namun intensitas peminjaman rendah (F=5,18 buku). Kategori favorit meliputi Pendidikan & Akademik, Ilmu Pengetahuan, dan Fiksi/Novel. Kehadiran fisik yang masih terjaga menjadikan kelompok ini target konversi paling potensial, di mana stimulus yang tepat untuk meningkatkan intensitas peminjaman berpotensi memberikan dampak paling cepat dan terukur.

3.3.3 Anggota Pasif (n=661, 46,2%)

Anggota Pasif memiliki R' tertinggi (0,7431) dengan rata-rata 808,54 hari ($\approx 2,2$ tahun) sejak kunjungan terakhir dan intensitas peminjaman terendah (F=5,15 buku). Temuan menarik pada kelompok ini adalah tingkat kepuasan tertinggi dibanding klaster lain, yakni 75,8% menyatakan Puas atau Sangat Puas, yang menyanggah anggapan bahwa kepasifan disebabkan ketidakpuasan layanan. Dengan demikian, strategi yang tepat adalah penjangkauan aktif (*outreach*) untuk mengurangi hambatan akses, bukan perbaikan kualitas layanan.

Tabel 4. Distribusi Tingkat Kepuasan Per *Cluster*

Label	Kurang Puas	Cukup	Puas	Sangat Puas	Total
Anggota Aktif	15 (7,5%)	36 (18,1%)	97 (48,7%)	51 (25,6%)	199
Anggota Sedang	37 (6,5%)	129 (22,6%)	270 (47,4%)	134 (23,5%)	570
Anggota Pasif	33 (5,0%)	127 (19,2%)	317 (48,0%)	184 (27,8%)	661

3.4 Analisis Kontribusi Dimensi RFM

Berdasarkan centroid hasil normalisasi (persamaan 1) pada Tabel 4, dimensi *Frequency* (F') terbukti menjadi pembeda utama antarklaster dengan selisih 0,5575 poin antara Anggota Aktif (F'=0,7305) dan dua klaster lainnya ($\pm 0,17$), jauh melampaui selisih R' (0,2971) maupun M' (0,048). Dimensi *Recency* (R') berkontribusi moderat dalam membedakan Anggota Sedang (R'=0,2460) dari Anggota Pasif (R'=0,7431), sementara *Monetary* (M') dengan rentang sangat kecil (0,2842–0,3322) tidak menghasilkan pembeda yang kuat sehingga variabel alternatif yang lebih diskriminatif perlu dieksplorasi pada penelitian selanjutnya. Secara keseluruhan, intensitas peminjaman (*Frequency*) lebih mencerminkan keterlibatan nyata anggota perpustakaan daerah dibandingkan dimensi R maupun M.

3.5 Rekomendasi Strategi Layanan

Berdasarkan profil karakteristik tiap *cluster*, dirumuskan tiga strategi layanan terdiferensiasi yang dapat diimplementasikan langsung tanpa modifikasi sistem:

- Strategi Retensi — Anggota Aktif:** Tingginya nilai *Frequency* (F'=0,7305) pada Anggota Aktif menunjukkan bahwa kelompok ini telah memiliki kebiasaan memanfaatkan perpustakaan secara intensif, sehingga strategi retensi lebih relevan dibandingkan promosi umum. Program loyalitas berbasis intensitas peminjaman seperti akses prioritas koleksi terbaru, undangan eksklusif ke diskusi buku, dan penunjukan sebagai "Duta Baca" dirancang untuk mempertahankan dan memperkuat kebiasaan yang sudah ada. Biaya mempertahankan anggota aktif secara inheren lebih rendah dari biaya reaktivasi, sehingga investasi pada kelompok ini memberikan *return* tertinggi per anggota [35].
- Strategi Peningkatan — Anggota Sedang:** Nilai R' terendah (0,2460) pada Anggota Sedang menunjukkan kehadiran fisik yang masih terjaga, sehingga hambatan utama bukan akses melainkan intensitas peminjaman. Program "Tantangan Baca Bulanan" dengan target bertahap, notifikasi proaktif koleksi terbaru sesuai kategori favorit, dan rekomendasi personal berbasis riwayat peminjaman dirancang untuk mengkonversi frekuensi kunjungan yang sudah ada menjadi intensitas peminjaman yang lebih tinggi [36]. Kelompok ini merupakan prioritas konversi karena potensi dampak paling cepat dan terukur [37].
- Strategi Reaktivasi — Anggota Pasif:** Kombinasi R' tertinggi (0,7431) dengan kepuasan layanan yang tinggi (75,8%) menunjukkan bahwa hambatan akses fisik menjadi faktor dominan kepasifan, bukan ketidakpuasan. Strategi penjangkauan aktif melalui pesan personal, program amnesti denda, pengembangan layanan digital (*e-catalog*,

peminjaman *online*) [38], dan perpustakaan keliling ke wilayah dengan konsentrasi Anggota Pasif tertinggi dirancang khusus untuk mengurangi hambatan akses tersebut bukan memperbaiki kualitas layanan yang sudah dinilai baik [39].

3.6 Perbandingan dengan Penelitian Sebelumnya

Temuan dalam penelitian ini sejalan dengan sejumlah kajian terdahulu sekaligus memperlihatkan perbedaan yang dapat dijelaskan secara kontekstual sekaligus menunjukkan perbedaan yang dapat dijelaskan secara kontekstual [9] yang menerapkan RFM dan K-Means pada data pelanggan komersial menghasilkan pola distribusi serupa segmen aktif minoritas, sedang, dan pasif mayoritas yang mengkonfirmasi validitas pendekatan ini pada konteks perpustakaan daerah. Perbedaan ditemukan pada dimensi pembeda utama: [10] menemukan *Monetary* sebagai pembeda dominan pada data ritel, sedangkan penelitian ini menemukan *Frequency* sebagai pembeda terkuat. Perbedaan ini wajar karena intensitas peminjaman lebih mencerminkan keterlibatan nyata anggota perpustakaan dibandingkan volume kategori. Dalam konteks perpustakaan, [11] menemukan bahwa pengguna dengan frekuensi tinggi cenderung memiliki preferensi kategori yang lebih spesifik, konsisten dengan profil Anggota Aktif dalam penelitian ini, sementara pendekatan [12] yang mengombinasikan K-Means dengan aturan asosiasi relevan sebagai arah pengembangan lanjutan untuk menghasilkan rekomendasi koleksi berbasis segmen yang telah terbentuk.

Untuk memastikan efektivitas implementasi ketiga strategi tersebut dapat dipantau secara terukur, setiap segmen dilengkapi dengan *Key Performance Indicator* (KPI) yang ditetapkan berdasarkan profil *cluster* yang terbentuk dan mengacu pada kerangka pengukuran kinerja layanan berbasis segmentasi [38][39][10]. Sebagaimana disajikan pada Tabel 5.

Tabel 5. KPI Strategi Layanan Per Segmen

Segmen	Strategi	KPI	Target
Anggota Aktif	Retensi	Retensi anggota	≥ 90%/tahun
		Rata-rata peminjaman	≥ 18 buku/kunjungan
Anggota Sedang	Peningkatan	Frekuensi peminjaman	≥ 50%/6 bulan
		Konversi ke aktif	≥ 15%/tahun
Anggota Pasif	Reaktivasi	Reaktivasi anggota	≥ 20%/12 bulan
		Rata-rata recency	≤ 400 hari/tahun

4. KESIMPULAN

Melalui penelitian ini, integrasi model RFM dengan algoritma K-Means berhasil diwujudkan sebagai kerangka segmentasi berbasis data yang diterapkan pada seluruh 1.430 anggota terdaftar Dinas Kearsipan dan Perpustakaan Daerah Kuala Kapuas. Pemodelan RFM berhasil merepresentasikan perilaku anggota melalui tiga dimensi, yaitu *Recency* (R), *Frequency* (F), dan *Monetary* (M), dengan dimensi *Monetary* diadaptasi sebagai volume peminjaman lintas kategori koleksi sehingga dapat diterapkan tanpa memerlukan atribut demografis tambahan. Selanjutnya, Algoritma K-Means berhasil membentuk tiga klaster anggota secara optimal yang dikonfirmasi melalui beberapa metrik evaluasi dengan hasil yang menunjukkan kualitas klasterisasi yang baik. Selain itu, dimensi *Frequency* terbukti menjadi pembeda utama antar klaster. Tiga klaster yang terbentuk terdiri atas Anggota Aktif (13,9%), Anggota Sedang (39,9%), dan Anggota Pasif (46,2%), yang masing-masing menghasilkan rekomendasi strategi layanan berupa retensi, peningkatan keterlibatan, dan reaktivasi. Temuan penelitian juga menunjukkan bahwa anggota pasif memiliki tingkat kepuasan yang relatif tinggi, sehingga rendahnya aktivitas kunjungan diduga lebih dipengaruhi oleh hambatan akses daripada kualitas layanan. Oleh karena itu, strategi reaktivasi melalui penjangkauan aktif dinilai lebih tepat dibandingkan hanya berfokus pada peningkatan kualitas layanan. Dengan demikian, secara keseluruhan, sinergi antara model RFM dan algoritma K-Means terbukti mampu menyajikan fondasi pengambilan keputusan yang berbasis data dan dapat diandalkan oleh pengelola perpustakaan untuk merancang intervensi layanan yang lebih personal dan efektif, semuanya tanpa menuntut perubahan pada sistem yang sudah ada.

UCAPAN TERIMAKASIH

Penulis menyampaikan apresiasi kepada Dinas Kearsipan dan Perpustakaan Kabupaten Kuala Kapuas yang telah memberikan izin dan dukungan data dalam pelaksanaan penelitian ini. Apresiasi yang sebesar-besarnya juga diberikan kepada Dosen Pembimbing yang telah meluangkan waktu untuk memberikan masukan dan arahnya, kepada kedua orang tua atas doa yang selalu dipanjatkan serta dukungan yang tak pernah berhenti, dan kepada teman-teman yang senantiasa memberikan semangat selama proses penelitian ini berlangsung.

REFERENCES

- [1] Y. Wei, I. K. D. Machica, J. C. T. Arroyo, M. L. P. Sabayle, and A. J. P. Delima, "Analysis of students' borrowing behavior in local university library using data mining technique," *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, no. 7, pp. 55–63, 2022, doi: 10.46338/ijetae0722_07.
- [2] D. Sundariyati, "Inovasi Layanan Perpustakaan Berbasis Inklusi Sosial di Dinas Kearsipan dan Perpustakaan Kabupaten Kebumen," *Pustablibia: Journal of Library and Information Science*, vol. 8, no. 1, pp. 19–43, 2024, doi: 10.18326/pustablibia.v8i1.19-43.
- [3] Sutrisno, "Implementasi Pengambilan Keputusan Berbasis Data oleh Kepala Sekolah dalam Peningkatan Mutu Pendidikan," *Jurnal Administrasi dan Manajemen Pendidikan*, vol. 5, no. 1, pp. 56–67, 2022.
- [4] A. Pujiyanto, "Big Data Dalam Pengambilan Keputusan Organisasi," *COSMIC: Journal of Computer Science and Information Technology*, 2023, [Online]. Available: <https://journal.aira.or.id/index.php/cosmic>
- [5] F. M. Amin, E. F. Rusydiyah, and A. N. Azizah, "Personalized library book recommendations using K-Means clustering and association rules," *Journal of Scientometric Research*, vol. 14, no. 1, pp. 32–45, 2025, doi: 10.5530/jscires.20251005.
- [6] P. Anitha and M. M. Patil, "RFM model for customer purchase behavior using K-Means algorithm," *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 5, pp. 1785–1792, 2022, doi: 10.1016/j.jksuci.2019.12.011.
- [7] E. S. Raharjo and T. Sugiat, "Analysis of recency frequency monetary (RFM) using Python and Matplotlib in optimizing marketing strategies at Dhiza Petshop retail store in Tanjung Selor," *Dinasti International Journal of Education Management and Social Science*, vol. 6, no. 5, pp. 4086–4110, 2025, doi: 10.38035/dijemss.v6i5.4679.
- [8] C. G. Wong, T. Nguyen, and K. Lee, "Application of Hierarchical and K-Means Clustering in E-Commerce Customer Segmentation," *International Journal of Information Management Data Insights*, vol. 2, no. 1, p. 100055, 2022, doi: 10.1016/j.jjime.2021.100055.
- [9] O. N. Akande, E. O. Asani, and B. T. Dautare, "Customer segmentation through RFM analysis and K-means clustering: Leveraging data-driven insights for effective marketing strategy," *Ceddi Journal of Information System and Technology*, vol. 3, no. 1, pp. 14–25, 2024, doi: 10.56134/jst.v3i1.81.
- [10] J. S. Aqil, M. K. Anam, and A. R. S. Sianipar, "Evaluating retail customer loyalty using RFM model and K-Means clustering algorithm," *Procedia Comput. Sci.*, vol. 227, pp. 344–353, 2023, doi: 10.1016/j.procs.2023.10.533.
- [11] R. A. Putri, D. A. Nurhasanah, and M. I. Perangin-Angin, "Rekomendasi pemilihan bahan bacaan pengunjung perpustakaan menggunakan metode K-Means clustering dan aturan asosiasi," *Komputika: Jurnal Sistem Komputer*, vol. 14, no. 2, pp. 125–132, 2025, doi: 10.34010/komputika.v14i1.13298.
- [12] A. F. Zabidi, "Penerapan algoritma K-Means untuk pengelompokan koleksi perpustakaan dengan data mining," *Media Jurnal Informatika*, vol. 16, no. 2, p. 233, 2024, doi: 10.35194/mji.v16i2.4814.
- [13] H. Hairani, D. Susilowati, I. P. Lestari, K. Marzuki, and L. Z. A. Mardedi, "Segmentasi Lokasi Promosi Penerimaan Mahasiswa Baru Menggunakan Metode RFM dan K-Means Clustering," *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 2, pp. 275–282, 2022, doi: 10.30812/matrik.v21i2.1542.
- [14] D. A. Fakhri, S. Defit, and Sumijan, "Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering," *Jurnal Informasi dan Teknologi*, vol. 3, no. 3, pp. 160–166, 2021, doi: 10.37034/jidt.v3i3.137.
- [15] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-Means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci. (N. Y.)*, vol. 622, pp. 178–210, 2023, doi: 10.1016/j.ins.2022.11.139.
- [16] A. Handojo, N. Pujawan, B. Santosa, and M. L. Singgih, "A multi layer recency frequency monetary method for customer priority segmentation in online transaction," *Cogent Eng.*, vol. 10, no. 1, p. 2162679, 2023, doi: 10.1080/23311916.2022.2162679.
- [17] B. N. Yuliasih, Herman, Sunardi, and H. Yuliansyah, "Evaluation of K-Means clustering using silhouette score method on customer segmentation," *ILKOM Jurnal Ilmiah*, vol. 16, no. 3, pp. 330–342, 2024, doi: 10.33096/ilkom.v16i3.2325.
- [18] M. Jordy, A. Triayudi, and I. D. Sholihati, "Analisis segmentasi recency dan customer value pada AVANA Indonesia dengan algoritma K-Means dan model RFM (Recency, Frequency and Monetary)," *Journal of Information System Research*, vol. 4, no. 2, pp. 579–589, 2023, doi: 10.47065/josh.v4i2.2950.

- [19] M. S. Uddin, M. A. Kabir, and S. Islam, "Applying RFM model and clustering techniques for analyzing user behavior and borrowing patterns in academic libraries," *International Journal of Information Management Data Insights*, vol. 3, no. 1, pp. 100–112, 2023, doi: 10.1016/j.jjimei.2023.100152.
- [20] C. Wongoutong, "The impact of neglecting feature scaling in K-means clustering," *PLoS One*, vol. 19, no. 10, p. e0310839, 2024, doi: 10.1371/journal.pone.0310839.
- [21] M. Aliyev, E. Ahmadov, and H. Gadirli, "Optimizing Recency, Frequency, and Monetary (RFM) Dimensions via K-Means Normalization Techniques for User Behavior Profiling," *International Journal of Computer and Information Engineering*, vol. 16, no. 4, pp. 182–191, 2022, doi: 10.1155/2022/8884227.
- [22] W. A. Silamantha and K. Hadiono, "Analisis RFM dan K-Means Clustering untuk Segmentasi Pelanggan pada PT. Sanutama Bumi Arto," *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, vol. 5, no. 3, pp. 1297–1305, 2024, doi: 10.30645/kesatria.v5i3.448.
- [23] G. Aslantaş, M. Gençgöl, M. Rumelli, M. Özsarac, and G. Bakırlı, "Customer Segmentation Using K-Means Clustering Algorithm and RFM Model," *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, vol. 25, no. 74, pp. 491–503, May 2023, doi: 10.21205/deufmd.2023257418.
- [24] D. S. Nugraha, I. Thoib, N. Sururi, F. Bayu, and B. P. Candra, "Segmentasi Pelanggan Berbasis Analisis RFM Menggunakan Algoritma K-Means Clustering," *Jurnal Komputer Teknologi Informasi Sistem Komputer (JUKTISI)*, vol. 4, no. 2, pp. 1361–1369, 2025, doi: 10.62712/juktisi.v4i2.641.
- [25] L. Y. Winarto, S. C. Kumajas, and A. Hasibuan, "Penerapan algoritma K-Means untuk mengelompokan daerah asal pada siswa SMK Kristen Harapan Rantepao Toraja Utara," *Jurnal Minfo Polgan*, vol. 14, no. 2, pp. 2318–2328, 2025, doi: 10.33395/jmp.v14i2.15431.
- [26] D. Anggreani, Nurmisa, D. Setiawan, and Lukman, "Optimization of K-Means clustering method by using elbow method in predicting blood requirement of Pelamonia Hospital Makassar," *Internet of Things and Artificial Intelligence Journal*, vol. 4, no. 3, pp. 541–550, 2024, doi: 10.31763/iota.v4i3.755.
- [27] E. U. Oti, M. O. Olusola, F. C. Eze, and S. U. Enogwe, "Comprehensive review of K-Means clustering algorithms," *International Journal of Advances in Scientific Research and Engineering*, vol. 7, no. 8, pp. 64–69, 2021, doi: 10.31695/IJASRE.2021.34050.
- [28] R. Ardiansyah, N. Trezandy, I. Skandar, M. Ilman, and S. Sahril, "Recency, frequency, and monetary-based customer segmentation using K-Means for analysing transactional behaviour in a service-based MSME," *Green Intelligent Systems and Applications*, vol. 6, no. 1, pp. 65–80, 2026, doi: 10.53623/gisa.v6i1.919.
- [29] J. Ahmed, S. M. Islam, and M. A. Rahman, "Customer value segmentation based on RFM model and K-Means clustering algorithm," *Int. J. Data Sci. Anal.*, vol. 14, no. 3, pp. 245–255, 2023, doi: 10.1007/s41060-022-00401-5.
- [30] R. A. Sukanto, M. R. Priyadi, and D. Purwanti, "Segmentasi pelanggan menggunakan model RFM dan algoritma K-Means berbasis nilai centroid," *Jurnal Edukasi dan Penelitian Informatika*, vol. 9, no. 1, pp. 88–95, 2023, doi: 10.26418/jp.v9i1.59412.
- [31] N. Nurmilayanti and Hartono, "Implementasi Metode Clustering Sebagai Penunjang Strategi dalam Manajemen Pelanggan," *Jurnal Fasilkom*, vol. 13, no. 3, pp. 605–613, 2023, doi: 10.37859/jf.v13i3.5617.
- [32] L. Chen and Y. Wang, "Evaluating the Effectiveness of AI-Powered Innovations in Public Libraries," *Libr. Inf. Sci. Res.*, vol. 45, no. 2, p. 101192, 2023, doi: 10.1016/j.lisr.2023.101192.
- [33] Pratiwi, "Penerapan Pendekatan Klasterisasi pada Data Keanggotaan untuk Menentukan Lokasi Sosialisasi Literasi yang Strategis," *Jurnal Muara Sains, Teknologi, Kedokteran dan Ilmu Kesehatan*, vol. 8, no. 1, pp. 1–10, 2024, doi: 10.24912/jmstkik.v8i1.26766.
- [34] J. C. Cheng, "Enhancing User Retention and Engagement through Personalized Loyalty Programs and Exclusive Events in Academic and Public Libraries," *J. Libr. Adm.*, vol. 63, no. 4, pp. 432–451, 2023, doi: 10.1080/01930826.2023.2198765.
- [35] R. D. Ananda, H. Santoso, and F. Firdaus, "Analisis strategi reaktivasi anggota pasif berdasarkan segmentasi RFM menggunakan K-Means clustering," *Jurnal Teknologi dan Sistem Informasi*, vol. 5, no. 1, pp. 62–70, 2024, doi: 10.33330/jtsi.v5i1.2917.
- [36] K. L. Chin and M. N. Richards, "Digital platform user retention strategies and proactive communication models for dormant user reactivation," *Front. Educ. (Lausanne)*, vol. 9, p. 1385278, 2024, doi: 10.3389/educ.2024.1385278.
- [37] S. Zhang, Y. Wang, and L. International, "Customer retention strategies and KPI optimization based on RFM segmentation and data-driven clustering models," *Journal of Retailing and Consumer Services*, vol. 71, p. 103214, 2023, doi: 10.1016/j.jretconser.2022.103214.
- [38] X. Li, T. L. Olsen, and S. M. Gilbert, "Optimizing user segment migration and behavioral elevation strategies using data-driven RFM modeling," *Operational Research*, vol. 23, no. 2, pp. 1–22, 2023, doi: 10.1007/s12351-023-00772-5.
- [39] H. Hwang, T. Jung, and E. Kim, "An empirical study on proactive reactivation strategies for dormant users based on behavioral RFM segmentation," *Electron. Commer. Res. Appl.*, vol. 58, p. 101243, 2023, doi: 10.1016/j.elerap.2023.101243.