

Analisis Pengelompokan Pengunjung Perpustakaan Fakultas Sains dan Teknologi Menggunakan Metode K-Means Clustering

Analysis of Visitor Grouping at the Faculty of Science and Technology Library Using the K-Means Clustering Method.

Maha Valne Datin^{*,1}, Muhammad Iqbal²

¹Magister Teknologi Informasi, Universitas Pembangunan Panca Budi, Sumatera Utara
Indonesia

² Magister Teknologi Informasi, Universitas Pembangunan Panca Budi, Sumatera Utara
Indonesia

*Penulis Korespondensi
Email: mhvlndtn04@gmail.com

Abstrak. Penelitian ini bertujuan untuk mengelompokkan pengunjung Perpustakaan Fakultas Sains dan Teknologi berdasarkan jurusan dan intensitas kunjungan menggunakan metode K-Means Clustering, guna mendukung peningkatan kualitas layanan perpustakaan secara tepat sasaran. Data yang digunakan merupakan data absensi pengunjung Perpustakaan Fakultas Sains dan Teknologi tahun 2025 yang mencakup informasi program studi dan frekuensi kunjungan mahasiswa. Metode yang digunakan dalam penelitian ini adalah K-Means Clustering, yaitu salah satu teknik data mining yang berfungsi untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik. Sebelum dilakukan proses clustering, data melalui tahap pra-pemrosesan yang meliputi pembersihan data, seleksi fitur, normalisasi nilai, serta konversi data kategorik ke dalam bentuk numerik. Jumlah klaster yang digunakan dalam penelitian ini adalah tiga klaster yang merepresentasikan tingkat intensitas kunjungan rendah, sedang, dan tinggi. Hasil penelitian menunjukkan bahwa metode K-Means Clustering mampu mengelompokkan pengunjung perpustakaan secara objektif dan sistematis dengan distribusi klaster sebesar 20,6% pada klaster intensitas rendah (C1), 54,4% pada klaster intensitas sedang (C2), dan 25,0% pada klaster intensitas tinggi (C3), serta mengungkap adanya perbedaan pola pemanfaatan perpustakaan berdasarkan jurusan dan intensitas kunjungan mahasiswa. Hasil clustering ini dapat dimanfaatkan sebagai dasar pengambilan keputusan dalam pengembangan layanan, penyesuaian koleksi, serta peningkatan efektivitas pengelolaan perpustakaan.

Kata kunci: K-Means Clustering, Data Mining, Perpustakaan, Intensitas Kunjungan, Pengelompokan Pengunjung

Abstract. This study aims to classify visitors of the Faculty of Science and Technology Library based on their academic programs and visit intensity in order to support the targeted improvement of library service quality. The data used in this study consist of visitor attendance records of the Faculty of Science and Technology Library in 2025, which include information on study programs and the frequency of student visits. The method applied in this research is K-Means Clustering, a data mining technique used to group data based on the similarity of their characteristics. Prior to the clustering process, the data underwent a preprocessing stage, including the selection of relevant variables and the conversion of categorical data into numerical form. The number of clusters used in this study was three, representing low, medium, and high levels of visit intensity. The results show that the K-Means Clustering method grouped library visitors objectively and

systematically, with a cluster distribution of 20.6% in the low-intensity cluster (C1), 54.4% in the medium-intensity cluster (C2), and 25.0% in the high-intensity cluster (C3), revealing differences in library utilization patterns based on academic programs and students' visit intensity. These clustering results can be used as a basis for decision-making in service development, collection adjustment, and improving the effectiveness of library management.

Keywords: K-Means Clustering, Data Mining, Library, Visit Intensity, Visitor Grouping.

1. Pendahuluan

Perpustakaan perguruan tinggi memiliki peran strategis dalam mendukung pelaksanaan tridharma perguruan tinggi, khususnya dalam penyediaan sumber informasi bagi mahasiswa dan dosen (Wardani et al., 2023). Tingkat pemanfaatan perpustakaan dapat mencerminkan kebutuhan informasi sivitas akademika serta efektivitas layanan yang disediakan (Fadilah et al., 2023). Perpustakaan tentunya menyediakan berbagai macam koleksi akademik sampai dengan non akademik yang dapat dimanfaatkan oleh pemustaka (Rosiana et al., 2025). Berdasarkan data absensi pengunjung Perpustakaan FST tahun 2025, tercatat bahwa pengunjung berasal dari berbagai program studi dan jenjang semester yang berbeda. Data tersebut mencakup informasi jenis kelamin, program studi, serta semester mahasiswa. Keberagaman karakteristik pengunjung ini menunjukkan adanya potensi perbedaan pola dan kebutuhan dalam pemanfaatan layanan perpustakaan, yang apabila tidak dianalisis lebih lanjut dapat menyulitkan pihak pengelola dalam melakukan evaluasi dan perencanaan layanan secara optimal (Abdussalam Amrullah et al., 2022).

Klasterisasi adalah proses pengelompokan data serupa ke dalam kelompok yang berbeda, atau lebih tepatnya partisi dari sebuah dataset ke dalam subset, sehingga data dalam setiap subset memiliki makna yang bermanfaat (Safitri, 2024). Dalam konteks perpustakaan, klasterisasi digunakan untuk mengelompokkan pengunjung berdasarkan kesamaan karakteristik seperti jurusan, semester, dan intensitas kunjungan, sehingga pihak pengelola dapat memahami pola pemanfaatan layanan secara lebih terstruktur dan menjadikannya dasar dalam pengembangan fasilitas serta penyesuaian koleksi. Metode K-Means Clustering digunakan untuk mengidentifikasi pola dan segmentasi pengunjung berdasarkan kemiripan karakteristik, sehingga hasil pengelompokan dapat dimanfaatkan dalam pengembangan layanan dan penyesuaian koleksi perpustakaan sesuai kebutuhan akademik masing-masing klaster (Dista & Abdulloh, 2022).

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengelompokkan pengunjung Perpustakaan Fakultas Sains dan Teknologi berdasarkan jurusan dan intensitas kunjungan menggunakan metode K-Means Clustering, guna mengidentifikasi pola pemanfaatan perpustakaan dan menjadikannya dasar peningkatan layanan, fasilitas, serta penyesuaian koleksi secara tepat sasaran. Hasil pengelompokan diharapkan dapat memberikan gambaran yang lebih jelas mengenai segmentasi pengunjung Perpustakaan FST, sehingga dapat menjadi dasar dalam pengambilan keputusan, seperti pengembangan koleksi, peningkatan fasilitas, serta penyusunan strategi layanan yang lebih tepat sasaran sesuai dengan karakteristik mahasiswa (Marcelina & Kurnia, 2023).

2. Metode

Algoritma K-Means merupakan salah satu algoritma clustering bertipe partitional yang bekerja dengan cara membagi dataset ke dalam sejumlah kelompok berdasarkan kemiripan karakteristik data. Metode K-Means dipilih dalam penelitian ini karena memiliki keunggulan

dibandingkan metode clustering lainnya, yaitu proses komputasi yang cepat dan efisien, kemudahan implementasi, serta kemampuan menangani dataset berukuran sedang hingga besar secara sistematis. Selain itu, K-Means didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai centroid awalnya, sehingga proses pengelompokan dapat dilakukan secara terstruktur dan hasilnya mudah diinterpretasikan (Wisanta & Marlim, 2021). Algoritma K-Means menggunakan proses secara berulang-ulang untuk mendapatkan basis data cluster hingga posisi centroid tidak lagi mengalami perubahan yang signifikan (Putriani et al., 2023).

2.1 Jenis dan Sumber Data

Penelitian ini menggunakan data sekunder berupa data absensi pengunjung Perpustakaan Fakultas Sains dan Teknologi tahun 2025 dengan jumlah keseluruhan sebanyak 68 data. Data yang digunakan mencakup informasi program studi, semester, jenis kelamin, dan intensitas kunjungan mahasiswa. Data tersebut digunakan sebagai dasar untuk melakukan pengelompokan pengunjung perpustakaan menggunakan metode K-Means Clustering.

Tabel 1. Data Pengunjung Perpustakaan Fak. Sains dan Teknologi.

No	Nama	Jurusan	Semester	JK
1	Sabrina Nasution	3	7	2
2	Ade Irma	5	7	2
3	Annisa Salsabila	5	7	2
4	Della Damayanti	5	7	2
...	...	...	...	...
65	Mesyadi	5	8	1
66	Barratun Nisa Hasra	4	8	2
67	Mutiara Alya Utami	4	8	2
68	Muhammad Saprin	1	5	1

Keterangan:

Jurusan Ilmu Komputer Nilainya	: 1	Semester 9 s.d 14 Nilainya	: 9
Jurusan Sistem Informasi Nilainya	: 2	JK	: Jenis Kelamin
Jurusan Matematika Nilainya	: 3	Jenis Kelamin Laki-laki Nilainya	: 1
Jurusan Biologi Nilainya	: 4	Jenis Kelamin Perempuan Nilainya	: 2
Jurusan Fisika Nilainya	: 5		

2.2 Metode K-Means

Metode K-Means digunakan untuk mengelompokkan data ke dalam sejumlah klaster berdasarkan tingkat kemiripan karakteristik data. Dalam penelitian ini, jumlah klaster ditentukan sebanyak tiga klaster ($K = 3$) (Alif et al., 2022). Algoritma K-Means memiliki performa yang baik dari algoritma lainnya, baik dari sisi nilai average within centroid distance dan kompleksitas waktu. Algoritma K-Means menggunakan centroid (rata-rata) sebagai model dari cluster (Dana & Ajiz, 2022). Langkah-langkah algoritma K-Means adalah sebagai berikut:

1. Menentukan jumlah klaster (K).
2. Menentukan pusat klaster (centroid) awal secara acak.
3. Menghitung jarak setiap data ke masing-masing centroid.
4. Mengelompokkan data ke klaster dengan jarak terdekat.
5. Memperbarui centroid berdasarkan rata-rata data dalam setiap klaster.
6. Mengulangi langkah 3–5 hingga centroid tidak mengalami perubahan.

2.3 Perhitungan Jarak (*Euclidean Distance*)

Metode K-Means mengarahkan pengelompokan data ke dalam beberapa kelompok, di mana data dalam satu kelompok memiliki karakteristik yang sama dan berbeda karakteristik dengan data di kelompok yang lain (L Maukar et al., 2022). Perhitungan jarak antar data dengan centroid dilakukan menggunakan rumus Euclidean Distance sebagai berikut (Sonia & Putri, 2023):

$$(ai, bj) = \sqrt{\sum(ai - cj)^2} \quad (1)$$

Keterangan:

ai : Data Kriteria,

bj : *Centroid* pada *Cluster* ke-*j*

3. Hasil dan Pembahasan

Penggunaan K-Means untuk menganalisis pola kunjungan pengunjung Perpustakaan Fakultas Sains dan Teknologi berfungsi mengidentifikasi kelompok pengunjung yang memiliki karakteristik serupa, seperti jurusan, semester, dan intensitas kunjungan mahasiswa, sehingga hasil pengelompokan dapat dimanfaatkan sebagai dasar peningkatan layanan perpustakaan secara tepat sasaran (Rosiana et al., 2025).

3.1 Inisialisasi

Pada tahap awal, dipilih K data secara acak sebagai centroid awal (Lubis et al., 2025). Dalam penelitian ini, centroid awal ditentukan berdasarkan data ke-7, ke-8, dan ke-9

Tabel 2. Centroid Awal

Cluster	Jurusan	Semester	JK
C1 (Rendah)	5	7	2
C2 (Sedang)	1	5	1
C3 (Tinggi)	4	9	2

3.2 Proses Iterasi K-Means

Proses iterasi dilakukan untuk memperbarui posisi centroid secara berulang hingga tidak terjadi perubahan signifikan. Berikut adalah ringkasan hasil setiap iterasi beserta interpretasinya.

Iterasi 1

Perhitungan jarak antara setiap data pengunjung dan centroid awal dilakukan pada tahap ini. Hasil perhitungan tersebut menjadi dasar dalam menentukan cluster setiap data menggunakan rumus berikut:

$$(ai, bj) = \sqrt{\sum(ai - cj)^2} \quad (2)$$

Jarak centroid dari data pertama dengan centroid 1,2,3

$$C1 = (ai, bj) = \sqrt{\sum(3 - 5)^2 + (7 - 7)^2 + (2 - 2)^2} = 2$$

$$C2 = (ai, bj) = \sqrt{\sum(3 - 1)^2 + (7 - 5)^2 + (2 - 1)^2} = 2$$

$$C3 = (ai, bj) = \sqrt{\sum(3 - 4)^2 + (7 - 9)^2 + (2 - 2)^2} = 2,236067977$$

Berdasarkan hasil perhitungan jarak data pengunjung, diperoleh nilai jarak untuk setiap data terhadap pusat cluster yang telah ditentukan. Langkah selanjutnya dapat dilihat pada tabel di bawah ini:

Tabel 3. Hasil Perhitungan Data Pengunjung Dengan Centroid Awal

No	Nama	C1	C2	C3	Cluster
1	sabrina nasution	2	3	2.236067977	C1
2	Ade Irma	0	4.582575695	2.236067977	C1
3	Annisa salsabila	0	4.582575695	2.236067977	C1
4	Della Damayanti	0	4.582575695	2.236067977	C1
...	...	...	...	...	...
65	Mesyadi	1.414213562	5	1.732050808	C1
66	Barratun Nisa Hasra	1.414213562	4.358898944	1	C3
67	Mutiara alya utami	1.414213562	4.358898944	1	C3
68	Muhammad Saprin	4.582575695	0	5.099019514	C2

Setelah mengetahui hasil clustering pada iterasi awal, langkah selanjutnya adalah menghitung rata-rata data berdasarkan cluster yang sama. Rata-rata data dari cluster tersebut akan menjadi nilai centroid kedua. Berikut adalah hasil perhitungan Centroid baru.

Tabel 4. Perhitungan Centroid Baru Awal

No	Nama	C1			C2			C3		
		Jurusan	Semester	JK	Jurusan	Semester	JK	Jurusan	Semester	JK
1	Sabrina nasution	3	7	2	-	-	-	-	-	-
2	Ade Irma	5	7	2	-	-	-	-	-	-
3	Annisa salsabila	5	7	2	-	-	-	-	-	-
4	Della Damayanti	5	7	2	-	-	-	-	-	-
...	...	...	...	...	...	...	...	...	...	...
65	Mesyadi	5	8	1	-	-	-	-	-	-
66	Barratun Nisa Hasra	-	-	-	-	-	-	4	8	2
67	Mutiara alya utami	-	-	-	-	-	-	4	8	2
68	Muhammad Saprin	-	-	-	1	5	1	-	-	-
Jumlah		47	72	19	59	262	74	27	86	15
Rata-rata		4.7	7.2	1.9	1.229167	5.458333	1.541667	2.7	8.6	1.5

Keterangan:

C1, C2, C3 merupakan Klaster 1, Klaster 2, dan Klaster 3 hasil pengelompokan data pengunjung menggunakan metode K-Means pada iterasi tersebut. Jumlah adalah total nilai atribut (Jurusan, Semester, JK) dari seluruh anggota pada masing-masing klaster. Rata-rata merupakan nilai rata-rata atribut pada masing-masing klaster, yang digunakan sebagai nilai centroid baru untuk iterasi berikutnya. Sedangkan tanda “-” pada tabel menunjukkan bahwa data pengunjung tersebut bukan merupakan anggota klaster yang bersangkutan pada iterasi tersebut, sehingga nilai atributnya tidak disertakan dalam perhitungan jumlah dan rata-rata klaster tersebut.

Nilai centroid baru diperoleh berdasarkan hasil perhitungan yang ditunjukkan pada tabel sebelumnya. Hasil pembaruan centroid untuk masing-masing cluster dapat dilihat pada tabel berikut.

Tabel 5. Pusat Centroid Baru Iterasi 2

Cluster	Jurusan	Semester	JK
C1 (Rendah)	4.7	7.2	1.9
C2 (Sedang)	1.229167	5.458333	1.541666667
C3 (Tinggi)	2.7	8.6	1.5

Iterasi 2

Pada iterasi ke-2, dilakukan perhitungan jarak setiap data pengunjung terhadap centroid yang baru diambil dari tabel 5. Hasil perhitungan tersebut digunakan untuk menentukan kembali cluster yang sesuai bagi masing-masing data.

Tabel 6. Hasil Perhitungan Data Pengunjung Dengan Centroid Awal Iterasi 2

No	NAMA	C1	C2	C3	Cluster
1	Sabrina Nasution	1.71464282	2.392207401	1.702938637	C3
2	Ade Irma	0.374165739	4.099510896	2.846049894	C1
3	Annisa Salsabila	0.374165739	4.099510896	2.846049894	C1
4	Della Damayanti	0.374165739	4.099510896	2.846049894	C1
...	...	...	...	...	...
65	Mesyadi	1.240967365	4.579591275	2.42899156	C1
66	Barratun Nisa Hasra	1.067707825	3.787830019	1.516575089	C1
67	Mutiara Alya Utami	1.067707825	3.787830019	1.516575089	C1
68	Muhammad Saprin	4.397726685	0.74564709	4.01248053	C2

Setelah mengetahui hasil clustering pada iterasi 2, langkah selanjutnya adalah menghitung rata-rata data berdasarkan cluster yang sama. Rata-rata data dari cluster tersebut akan menjadi nilai centroid kedua.

Tabel 7. Perhitungan Centroid Baru Iterasi 2

No	Nama	C1			C2			C3		
		Jurusan	Semester	JK	Jurusan	Semester	JK	Jurusan	Semester	JK
1	Sabrina Nasution	-	-	-	-	-	-	3	7	2
2	Ade Irma	5	7	2	-	-	-	-	-	-
3	Annisa Salsabila	5	7	2	-	-	-	-	-	-
4	Della Damayanti	5	7	2	-	-	-	-	-	-
...	...	...	...	...	...	...	...	...	...	...
6	Mesyadi	5	8	1	-	-	-	-	-	-
5										
6	Barratun	4	8	2	-	-	-	-	-	-
6	Nisa Hasra									
6	Mutiara	4	8	2	-	-	-	-	-	-
7	Alya Utami									
6	Muhammad	-	-	-	1	5	1	-	-	-
8	Saprin									
Jumlah		56	89	22	59	262	74	18	69	12
Rata-Rata		4.66667	7.41667	1.833	1.229167	5.458333	1.54167	2.25	8.625	1.5

Keterangan:

C1, C2, C3 merupakan Klaster 1, Klaster 2, dan Klaster 3 hasil pengelompokan data pengunjung menggunakan metode K-Means pada iterasi tersebut. Jumlah adalah total nilai atribut (Jurusan, Semester, JK) dari seluruh anggota pada masing-masing klaster. Rata-rata adalah nilai rata-rata atribut pada masing-masing klaster, yang digunakan sebagai nilai centroid baru untuk iterasi berikutnya dan Tanda “-” pada tabel menunjukkan bahwa data pengunjung tersebut bukan merupakan anggota klaster yang bersangkutan pada iterasi tersebut, sehingga nilai atributnya tidak disertakan dalam perhitungan jumlah dan rata-rata klaster tersebut.

Hasil perhitungan pada tabel sebelumnya menghasilkan pembaruan nilai centroid pada masing-masing cluster. Nilai centroid baru dapat dilihat pada tabel berikut.

Tabel 8. Pusat Centroid Baru untuk Iterasi 3

Cluster	Jurusan	Semester	JK
C1 (Rendah)	4.666666667	7.416666667	1.833333333
C2 (Sedang)	1.229166667	5.458333333	1.541666667
C3 (Tinggi)	2.25	8.625	1.5

Iterasi 3

Pada iterasi ke-3, dilakukan perhitungan jarak setiap data pengunjung terhadap centroid yang diambil dari hasil tabel 8. Hasil perhitungan tersebut digunakan untuk menentukan kembali keanggotaan cluster masing-masing data.

Tabel 9. Hasil Perhitungan Data Pengunjung Dengan Centroid Awal Iterasi 3

No	NAMA	C1	C2	C3	CLUSTER
1	sabrina nasution	1.726026265	2.392207401	1.858258593	C1
2	Ade Irma	0.559016994	4.099510896	3.233129289	C1
3	Annisa salsabila	0.559016994	4.099510896	3.233129289	C1
4	Della Damayanti	0.559016994	4.099510896	3.233129289	C1
...	...	...	...	...	...
65	Mesyadi	1.070436048	4.579591275	2.864109809	C1
66	Barratun Nisa Hasra	0.901387819	3.787830019	1.92435054	C1
67	Mutiara alya utami	0.901387819	3.787830019	1.92435054	C1
68	Muhammad Saprin	4.469806111	0.74564709	3.866927075	C2

Setelah mengetahui hasil clustering pada iterasi 3, langkah selanjutnya adalah menghitung rata-rata data berdasarkan cluster yang sama. Rata-rata data dari cluster tersebut akan menjadi nilai centroid ketiga.

Tabel 10. Perhitungan Centroid Baru Iterasi 3

No	Nama	C1			C2			C3		
		Jurusan	Semester	JK	Jurusan	Semester	JK	Jurusan	Semester	JK
1	sabrina nasution	3	7	2	-	-	-	-	-	-
2	Ade Irma	5	7	2	-	-	-	-	-	-
3	Annisa salsabila	5	7	2	-	-	-	-	-	-

Analisis Pengelompokan Pengunjung Perpustakaan Fakultas Sains dan Teknologi Menggunakan Metode K-Means Clustering

No	Nama	C1			C2			C3		
		Jurusan	Semester	JK	Jurusan	Semester	JK	Jurusan	Semester	JK
4	Della Damayanti	5	7	2	-	-	-	-	-	-
...	...	...	...	...	...	...	...	...	...	...
6	Mesyadi	5	8	1	-	-	-	-	-	-
6	Barratun Nisa	4	8	2	-	-	-	-	-	-
6	Hasra Mutiara	4	8	2	-	-	-	-	-	-
7	alya utami	-	-	-	1	5	1	-	-	-
6	Muhamad Saprin	-	-	-	1	5	1	-	-	-
Jumlah		67	114	28	41	199	62	25	107	18
Rata-rata		4.4667	7.6	1.867	1.0513	5.1026	1.5897	1.7857	7.6429	1.2857

Keterangan:

C1, C2, C3 adalah Klaster 1, Klaster 2, dan Klaster 3 hasil pengelompokan data pengunjung menggunakan metode K-Means pada iterasi tersebut. Jumlah merupakan total nilai atribut (Jurusan, Semester, JK) dari seluruh anggota pada masing-masing klaster. Rata-rata adalah nilai rata-rata atribut pada masing-masing klaster, yang digunakan sebagai nilai centroid baru untuk iterasi berikutnya. Tanda “-” pada tabel menunjukkan bahwa data pengunjung tersebut bukan merupakan anggota klaster yang bersangkutan pada iterasi tersebut, sehingga nilai atributnya tidak disertakan dalam perhitungan jumlah dan rata-rata klaster tersebut.

Berdasarkan hasil perhitungan pada tabel di atas, diperoleh nilai centroid baru untuk masing-masing cluster. Nilai centroid tersebut disajikan pada tabel berikut.

Tabel 11. Pusat Centroid Baru Iterasi 4

Cluster	Jurusan	Semester	JK
C1 (Rendah)	4.466666667	7.6	1.866666667
C2 (Sedang)	1.051282051	5.102564103	1.58974359
C3 (Tinggi)	1.785714286	7.642857143	1.28571429

Iterasi 4

Pada iterasi ke-4, dilakukan perhitungan jarak setiap data pengunjung terhadap centroid yang baru. Hasil perhitungan tersebut digunakan untuk menentukan kembali keanggotaan cluster masing-masing data.

Tabel 12. Hasil Perhitungan Data Pengunjung Dengan Centroid Iterasi 3

No	Nama	C1	C2	C3	Cluster
1	Sabrina Nasution	1.590248059	2.750649914	1.548534528	C3
2	Ade Irma	0.813770374	4.400107583	3.354862447	C1
3	Annisa Salsabila	0.813770374	4.400107583	3.354862447	C1
4	Della Damayanti	0.813770374	4.400107583	3.354862447	C1
...	...	...	...	...	...
65	Mesyadi	1.093414631	4.933082781	3.246662336	C1
66	Barratun Nisa Hasra	0.628932075	4.154320961	2.353893865	C1
67	Mutiara Alya Utami	0.628932075	4.154320961	2.353893865	C1
68	Muhammad Saprin	4.419150245	0.600788437	2.771943987	C2

Setelah mengetahui hasil clustering pada iterasi 4, langkah selanjutnya adalah menghitung rata-rata data berdasarkan cluster yang sama. Rata-rata data dari cluster tersebut akan menjadi nilai centroid keempat.

Tabel 13. Perhitungan Centroid Baru Iterasi 4

No	Nama	C1			C2			C3		
		Jurusan	semester	JK	Jurusan	Semester	JK	Jurusan	Semester	JK
1	Sabrina Nasution	-	-	-	-	-	-	3	7	2
2	Ade Irma	5	7	2	-	-	-	-	-	-
3	Annisa Salsabila	5	7	2	-	-	-	-	-	-
4	Della Damayanti	5	7	2	-	-	-	-	-	-
6	Mesyadi	5	8	1	-	-	-	-	-	-
5	Barratun	4	8	2	-	-	-	-	-	-
6	Nisa Hasra	4	8	2	-	-	-	-	-	-
6	Mutiara	4	8	2	-	-	-	-	-	-
7	Alya Utami	-	-	-	1	5	1	-	-	-
6	Muhammad	-	-	-	1	5	1	-	-	-
8	Saprin	-	-	-	1	5	1	-	-	-
Jumlah		64	107	26	39	185	60	30	128	22
Rata-rata		4.5714	7.6429	1.8571	1.0541	5	1.6216	1.7647	7.5294	1.2941

Keterangan:

C1, C2, C3 adalah Klaster 1, Klaster 2, dan Klaster 3 hasil pengelompokan data pengunjung menggunakan metode K-Means pada iterasi tersebut. Jumlah merupakan total nilai atribut (Jurusan, Semester, JK) dari seluruh anggota pada masing-masing klaster. Rata-rata adalah nilai rata-rata atribut pada masing-masing klaster, yang digunakan sebagai nilai centroid baru untuk iterasi berikutnya. Tanda “-” pada tabel menunjukkan bahwa data pengunjung tersebut bukan merupakan anggota klaster yang bersangkutan pada iterasi tersebut, sehingga nilai atributnya tidak disertakan dalam perhitungan jumlah dan rata-rata klaster tersebut.

Hasil perhitungan pada iterasi ke-4 menghasilkan nilai centroid baru untuk masing-masing cluster. Nilai centroid tersebut dapat dilihat pada tabel berikut.

Tabel 14. Pusat Centroid Baru Iterasi 5

Cluster	Jurusan	Semester	JK
C1 (Rendah)	4.571428571	7.642857143	1.857142857
C2 (Sedang)	1.054054054	5	1.621621622
C3 (Tinggi)	1.764705882	7.529411765	1.294117647

Iterasi 5

Pada iterasi ke-5, dilakukan perhitungan jarak setiap data pengunjung terhadap centroid yang baru diambil dari tabel 14. Hasil perhitungan tersebut digunakan untuk menentukan kembali keanggotaan cluster masing-masing data.

Tabel 15. Hasil Perhitungan Data Pengunjung Dengan Centroid Awal Iterasi 4

No	Nama	C1	C2	C3	Cluster
1	Sabrina Nasution	1.703837206	2.816003519	1.5180574	C3
2	Ade Irma	0.785714286	4.440006712	3.353457133	C1
3	Annisa Salsabila	0.785714286	4.440006712	3.353457133	C1
4	Della Damayanti	0.785714286	4.440006712	3.353457133	C1
...	...	...	...	...	...

65	Mesyadi	1.022701505	4.995688426	3.282542689	C1
66	Barratun Nisa Hasra	0.688832197	4.221583555	2.390870757	C1
67	Mutiara Alya Utami	0.688832197	4.221583555	2.390870757	C1
68	Muhammad Saprin	4.524874561	0.623967372	2.658797501	C2

Setelah mengetahui hasil clustering pada iterasi 5, langkah selanjutnya adalah menghitung rata-rata data berdasarkan cluster yang sama. Rata-rata data dari cluster tersebut akan menjadi nilai centroid kelima.

Tabel 16. Perhitungan Centroid Baru Iterasi 5

No	Nama	C1			C2			C3		
		Jurusan	semester	JK	Jurusan	Semester	JK	Jurusan	Semester	JK
1	Sabrina Nasution	-	-	-	-	-	-	3	7	2
2	Ade Irma	5	7	2	-	-	-	-	-	-
3	Annisa Salsabila	5	7	2	-	-	-	-	-	-
4	Della Damayanti	5	7	2	-	-	-	-	-	-
...	...	...	...	...	...	...	...	...	...	...
65	Mesyadi	5	8	1	-	-	-	-	-	-
66	Barratun Nisa Hasra	4	8	2	-	-	-	-	-	-
67	Mutiara Alya Utami	4	8	2	-	-	-	-	-	-
68	Muhammad Saprin	-	-	-	1	5	1	-	-	-
Jumlah		64	107	26	39	185	60	30	128	22
Rata-rata		4.5714	7.6429	1.857	1.0541	5	1.621	1.7647	7.5294	1.294
				1			6			1

Keterangan:

C1, C2, C3 adalah Klaster 1, Klaster 2, dan Klaster 3 hasil pengelompokan data pengunjung menggunakan metode K-Means pada iterasi tersebut. Jumlah merupakan total nilai atribut (Jurusan, Semester, JK) dari seluruh anggota pada masing-masing klaster. Rata-rata merupakan nilai rata-rata atribut pada masing-masing klaster, yang digunakan sebagai nilai centroid baru untuk iterasi berikutnya. Tanda “-” pada tabel menunjukkan bahwa data pengunjung tersebut bukan merupakan anggota klaster yang bersangkutan pada iterasi tersebut, sehingga nilai atributnya tidak disertakan dalam perhitungan jumlah dan rata-rata klaster tersebut.

Posisi centroid tidak lagi berubah secara signifikan setelah beberapa iterasi, terutama karena data sudah cukup terstruktur dan pemilihan centroid awal dilakukan dengan baik. Hal ini menyebabkan algoritma K-Means berhenti sebelum mencapai batas iterasi maksimum yang ditentukan. Berdasarkan hasil perhitungan, dapat disimpulkan bahwa posisi klaster pada iterasi ke-4 tidak mengalami perubahan pada iterasi ke-5, sehingga algoritma dinyatakan konvergen dan proses iterasi dihentikan pada iterasi ke-5.

3.3 Hasil Clustering

Berdasarkan penerapan metode K-Means Clustering terhadap data absensi pengunjung Perpustakaan Fakultas Sains dan Teknologi tahun 2025, proses iterasi konvergen pada iterasi ke-5 dan menghasilkan tiga klaster pengunjung, yaitu C1, C2, dan C3, yang masing-masing merepresentasikan tingkat intensitas kunjungan rendah, sedang, dan tinggi, dengan proporsi anggota sebesar 20,6%, 54,4%, dan 25,0% dari total 68 pengunjung. Klaster C1 merepresentasikan kelompok pengunjung dengan intensitas kunjungan rendah. Mahasiswa pada klaster ini umumnya berada pada semester menengah hingga akhir, namun frekuensi kunjungan mereka ke

perpustakaan masih tergolong terbatas. Kondisi ini mengindikasikan bahwa pemanfaatan layanan perpustakaan oleh kelompok tersebut belum optimal, sehingga diperlukan upaya promosi yang lebih aktif, seperti sosialisasi layanan perpustakaan, program literasi informasi, maupun penjadwalan kunjungan wajib pada semester-semester tersebut guna mendorong peningkatan keterlibatan mahasiswa. Klaster C2 merupakan kelompok dengan intensitas kunjungan sedang dan menjadi klaster dengan jumlah anggota terbesar. Mayoritas mahasiswa pada klaster ini berada pada semester menengah dan memanfaatkan perpustakaan secara berkala sesuai kebutuhan akademik. Untuk mempertahankan sekaligus meningkatkan pola kunjungan tersebut, perpustakaan perlu memperkuat ketersediaan koleksi referensi yang relevan dengan mata kuliah pada semester menengah serta meningkatkan kenyamanan fasilitas agar frekuensi kunjungan kelompok ini dapat terus meningkat. Klaster C3 mencerminkan kelompok dengan intensitas kunjungan tinggi. Kelompok ini didominasi oleh mahasiswa semester akhir yang memiliki kebutuhan literatur lebih besar, sehingga menjadikan perpustakaan sebagai sarana utama dalam menunjang penyelesaian tugas akhir dan aktivitas akademik lainnya. Guna mendukung kebutuhan kelompok ini secara optimal, perpustakaan perlu memprioritaskan ketersediaan koleksi tugas akhir, jurnal ilmiah, serta menyediakan ruang baca yang kondusif.

Tabel 17. Hasil Akhir Klasterisasi

Klaster	Interpretasi Intensitas	Jumlah Anggota	Persentase	Centroid Jurusan	Centroid Semester	Centroid JK
C1	Rendah	14 mahasiswa	20,6%	4,57 ($\pm$ Biologi–Fisika)	7,64 (semester akhir)	1,86 (didominasi perempuan)
C2	Sedang	37 mahasiswa	54,4%	1,05 ($\pm$ Ilmu Komputer)	5,00 (semester menengah)	1,62 (perempuan & laki-laki relatif seimbang, condong perempuan)
C3	Tinggi	17 mahasiswa	25,0%	1,76 ($\pm$ Ilmu Komputer–Sistem Informasi)	7,53 (semester akhir)	1,29 (didominasi laki-laki)
Total	-	68 mahasiswa	100%	-	-	-

4. Kesimpulan

Berdasarkan hasil penerapan metode K-Means Clustering terhadap 68 data pengunjung Perpustakaan Fakultas Sains dan Teknologi tahun 2025, algoritma berhasil mengelompokkan pengunjung ke dalam tiga klaster yang konvergen pada iterasi ke-5. Klaster C1 (intensitas rendah) beranggotakan 14 mahasiswa (20,6%), klaster C2 (intensitas sedang) beranggotakan 37 mahasiswa (54,4%) sebagai kelompok terbesar, dan klaster C3 (intensitas tinggi) beranggotakan 17 mahasiswa (25,0%). Hasil ini menunjukkan bahwa mayoritas pengunjung berada pada tingkat pemanfaatan perpustakaan yang sedang, sementara hanya seperempat pengunjung yang tergolong sangat aktif memanfaatkan perpustakaan, terutama mahasiswa semester akhir dari jurusan Ilmu Komputer dan Sistem Informasi. Temuan ini mengindikasikan perlunya strategi berjenjang: penguatan promosi dan literasi informasi bagi klaster C1, penambahan koleksi referensi mata kuliah bagi klaster C2, serta prioritas pada koleksi tugas akhir dan jurnal ilmiah bagi klaster C3.

Secara keseluruhan, metode K-Means Clustering terbukti efektif dan efisien dalam mengelompokkan pengunjung perpustakaan secara objektif, sehingga dapat dijadikan dasar pengambilan keputusan berbasis data bagi pengelola perpustakaan dalam merancang layanan yang lebih tepat sasaran.

Daftar Pustaka

- Abdussalam Amrullah, Intam Purnamasari, Betha Nurina Sari, Garno, & Apriade Voutama. (2022). Analisis Cluster Faktor Penunjang Pendidikan Menggunakan Algoritma K-Means (Studi Kasus: Kabupaten Karawang). *Jurnal Informatika Dan Rekayasa Elektronik*, 5(2), 244–252. <https://doi.org/10.36595/jire.v5i2.701>
- Alif, D., Fadhillah, F., Faisol, A., Vendyansyah, N., Industri, F. T., Lahan, P., Kopi, T., Geografis, S. I., & Malang, K. (2022). *Penerapan Metode K-Means Clustering Pada Pemetaan Lahan Kopi Di Kabupaten Malang*. 6(1), 162–170.
- Dana, R. D., & Ajiz, A. (2022). *Clustering Data Persediaan Barang dengan Menggunakan Metode*. 7(1), 1–9.
- Dista, T. M., & Abdulloh, F. F. (2022). *Clustering Pengunjung Mall Menggunakan Metode K-Means dan Particle Swarm Optimization*. 6, 1339–1348. <https://doi.org/10.30865/mib.v6i3.4172>
- Fadilah, N., Ikhwan, A., & Alda, M. (2023). Pengembangan Sistem Informasi Perpustakaan Pada Dinas Perpustakaan dan Kearsipan Kota Medan Berbasis Android. *J-SISKO TECH (Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD)*, 6(2), 298–306. <https://doi.org/10.53513/jsk.v6i2.8228>
- L Maukar, A., Marisa, F., Widodo, A. A., & Nugraha, D. (2022). *Analisis data penerimaan mahasiswa baru berbasis k-means*. 6(2), 142–147.
- Lubis, B. O., Salim, A., Fibriany, F. W., Sidauruk, J., Informatika, S., Mining, D., & Bangunan, D. (2025). *Penerapan Metode K-Means Clustering Untuk Analisa Barang Terlaris*. 7(2), 624–638.
- Marcelina, D., & Kurnia, A. (2023). *Cluster Analysis of Small Medium Enterprise Performance with K-Means Clustering Algorithm Analisis Klaster Kinerja Usaha Kecil dan Menengah Menggunakan Algoritma K-Means Clustering*. 3(October), 293–301.
- Nurchayani, H. (2023). *Penelitian Strategi Pengembangan Koleksi Di Perpustakaan Pada Google Scholar : Sebuah Narrative Literature Review*. 10(1).
- Putriani, N., Yulianti, L., Alinse, R. T., & Dehasen, U. (2023). *Penerapan Metode Clustering K-Means Dalam Pengelompokkan Jumlah Pengunjung Pada*. 4307(3), 724–727.
- Rosiana, L., Yuadi, I., Airlangga, U., Timur, J., & Artikel, I. (2025). *K-Means Clustering untuk Analisis Tren Peminjaman Buku di Perpustakaan*. 7(1), 1–10.
- Safitri, E. (2024). Implementasi Algoritma K-Means Clustering Dalam Menentukan Strategi Marketing Dalam Penjualan Ikan (Studi Kasus: Grosir Ikan Tani Mas Tanjung Morawa). *JIKTEKS : Jurnal Ilmu Komputer Dan Teknologi Informasi*, 02(02), 23–33.
- Sonia, G., & Putri, R. A. (2023). Penerapan Metode K-Means Clustering Untuk Mengelompokkan Data Kelayakan Penerima Bantuan Renovasi Rumah. *Building of Informatics, Technology and Science (BITS)*, 5(2), 442–455. <https://doi.org/10.47065/bits.v5i2.4298>
- Wardani, F. K., Kurniawan, R., Informatika, T., Lunak, R. P., & K-means, A. (2023).

Pengelompokan Responsivitas Hasil Kuesioner Terhadap Penggunaan Perpustakaan Digital Menggunakan Algoritma K-Means. 203–212.

Wisanta, E. H., & Marlim, Y. N. (2021). Analisis Algoritma K-Means Untuk Clustering Kepuasan Pelayanan: Mall Pelayanan Publik Pekanbaru. *Seminar Nasional Informatika (SENATIKA) Prosiding SENATIKA 2021*, 223–228.