

## Penerapan Algoritma K-Means Clustering dalam Menganalisis Pola Peminjaman Buku di Perpustakaan

Rapel Aprilius Sigit<sup>1</sup>, Vitto Rezki Akbar<sup>2</sup>, Rahmaddeni<sup>3</sup>, Syarfi Aziz<sup>4</sup>, Ega Ranaldi Pebriansyah<sup>5</sup>

rafelas04@gmail.com<sup>1</sup>, vitorezkiakbar@gmail.com<sup>2</sup>, rahmaddeni@sar.ac.id<sup>3</sup>,

syarfi.aziz@gmail.com<sup>4</sup>, egaranaldi93@gmail.com<sup>5</sup>

<sup>1, 2, 3, 5</sup> Universitas Sains dan Teknologi Indonesia

<sup>4</sup> Institut Az Zuhra

---

### Informasi Artikel

Diterima : 15 Jul 2024  
Direview : 5 Agu 2024  
Disetujui : 1 Okt 2024

---

### Kata Kunci

*Data Mining, K-Means, Clustering, Python, Perpustakaan*

---

### Abstrak

Peminjaman buku di perpustakaan mencerminkan preferensi pembaca terhadap berbagai topik, memberikan wawasan untuk meningkatkan pengelolaan koleksi. Penelitian ini mengelompokkan buku dalam lima kategori: buku lama dengan halaman sedikit, buku modern dengan peminjaman singkat, buku klasik dengan durasi peminjaman lama, buku dengan halaman banyak dan peminjaman moderat, serta buku baru dengan variasi durasi peminjaman. Dataset dianalisis melalui pre-processing yang mencakup pembuatan fitur baru, normalisasi, dan standard scaling. Algoritma K-Means digunakan untuk mengelompokkan data berdasarkan jarak Euclidean, dengan evaluasi hasil menggunakan Davies-Bouldin Index (DBI), di mana nilai lebih rendah menunjukkan hasil yang lebih baik. Penentuan jumlah cluster optimal dilakukan dengan Metode Elbow, yang menunjukkan bahwa lima cluster adalah yang paling efektif. Penerapan K-Means Clustering menghasilkan lima cluster dengan distribusi informatif, dan evaluasi menggunakan DBI sebesar 0.50 menunjukkan kualitas clustering yang baik. Scatter plot menunjukkan distribusi cluster berdasarkan tahun terbit buku, jumlah halaman, dan durasi peminjaman, dari dataset yang terdiri dari 1323 catatan peminjaman.

---

### Keywords

*Data Mining, K-Means, Clustering, Python, Libraries*

---

### Abstract

*Library book borrowing reflects readers' preferences on various topics, providing insights for improving collection management. This study categorizes books into five main categories: old books with few pages, modern books with short borrowing periods, classic books with long borrowing durations, books with many pages and moderate borrowing, and new books with varying borrowing durations. The dataset is analyzed through pre-processing, including feature creation, normalization, and standard scaling. The K-Means algorithm is used to cluster data based on Euclidean distance, with results evaluated using the Davies-Bouldin Index (DBI), where a lower value indicates better clustering quality. The optimal number of clusters is determined using the Elbow Method, showing five clusters as the most effective. Applying K-Means Clustering produces five informative clusters, with a DBI of 0.50 indicating good clustering quality. Scatter plots illustrate cluster distribution based on publication year, number of pages, and borrowing duration, from a dataset of 1323 borrowing records.*

## A. Pendahuluan

Perpustakaan merupakan salah satu institusi penting dalam menjaga dan menyebarkan pengetahuan di masyarakat. Namun, manajemen koleksi buku dan pola peminjaman buku sering kali menjadi permasalahan yang kompleks bagi perpustakaan. Data mengenai pola peminjaman buku menjadi krusial untuk meningkatkan efisiensi pengelolaan koleksi buku serta memberikan pelayanan yang lebih baik kepada pengunjung[1]. Masalah ini menjadi sangat penting untuk diselesaikan karena akan membantu perpustakaan dalam mengoptimalkan layanan dan alokasi sumber daya. Pengelolaan koleksi yang tidak efisien dapat menyebabkan penurunan jumlah pengunjung, peningkatan biaya operasional, dan ketidakpuasan pengguna[2].

Peminjaman buku di perpustakaan sering kali mencerminkan preferensi dan minat pembaca terhadap berbagai topik. Analisis terhadap pola peminjaman ini dapat memberikan wawasan berharga yang dapat digunakan untuk meningkatkan pengelolaan koleksi perpustakaan[3]. Salah satu metode yang efektif untuk menganalisis data dalam jumlah besar dan kompleks adalah data mining, khususnya dengan algoritma K-Means Clustering. Metode ini memungkinkan pengelompokan data berdasarkan kesamaan tertentu, sehingga memudahkan dalam mengidentifikasi pola-pola tersembunyi[4]. K-Means Clustering dipilih karena kemampuannya dalam menangani data dalam jumlah besar dengan efisien dan memberikan hasil yang jelas serta terstruktur[5].

Dalam penelitian ini, peneliti akan mengelompokkan buku berdasarkan lima kategori utama: Buku yang diterbitkan lebih lama dengan jumlah halaman sedikit, Buku modern dengan peminjaman singkat, Buku klasik dengan durasi peminjaman lama, Buku dengan halaman banyak dan peminjaman moderat, dan Buku baru dengan variasi dalam durasi peminjaman. Klasifikasi ini akan membantu perpustakaan dalam memahami preferensi peminjam terhadap jenis buku tertentu dan durasi peminjaman yang terkait. Kriteria pemisahan ini didasarkan pada studi-studi sebelumnya yang menunjukkan bahwa durasi peminjaman dapat mencerminkan jenis dan relevansi buku terhadap peminjam.

Algoritma K-Means Clustering merupakan salah satu algoritma clustering yang paling populer dan sering digunakan dalam data mining. Algoritma ini bekerja dengan membagi data ke dalam beberapa cluster yang telah ditentukan sebelumnya, berdasarkan jarak Euclidean antar data[6]. Tujuan dari proses ini adalah untuk meminimalkan variasi di dalam cluster dan memaksimalkan variasi antar cluster. Dalam konteks peminjaman buku perpustakaan, K-Means Clustering dapat digunakan untuk mengelompokkan buku-buku berdasarkan pola peminjaman. Hasil pengelompokan ini dapat membantu perpustakaan dalam membuat keputusan strategis, seperti pengadaan buku baru, penataan koleksi, dan promosi buku[7].

Pemanfaatan K-Means Clustering dalam analisis pola peminjaman buku memiliki beberapa keuntungan. Pertama, metode ini dapat menangani data dalam jumlah besar dengan efisien. Kedua, hasil clustering dapat memberikan informasi yang jelas dan terstruktur tentang kelompok-kelompok buku yang memiliki pola peminjaman serupa. Ketiga, dengan identifikasi kelompok-kelompok ini, perpustakaan dapat lebih mudah memahami preferensi peminjam dan melakukan penyesuaian yang tepat untuk memenuhi kebutuhan mereka[8].

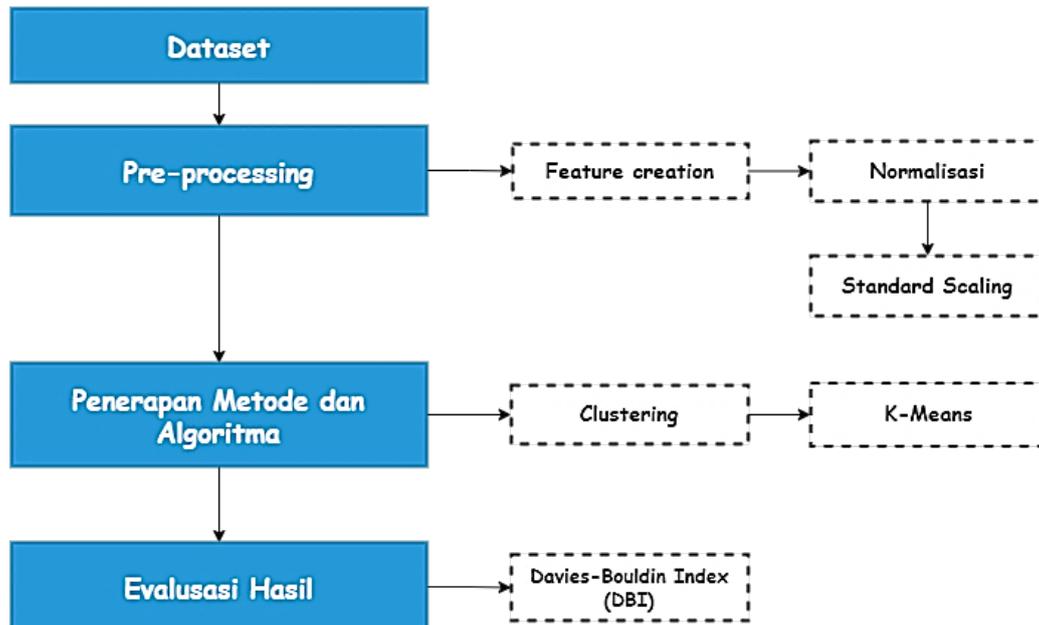
Penerapan data mining dengan menggunakan K-Means Clustering untuk pola peminjaman buku di perpustakaan juga dapat meningkatkan kualitas layanan perpustakaan. Sebagai contoh, dengan mengetahui kelompok buku yang sering dipinjam bersamaan, perpustakaan dapat menyusun rekomendasi buku yang lebih sesuai untuk pengguna. Selain itu, perpustakaan dapat mengatur ulang tata letak rak buku berdasarkan kelompok-kelompok yang telah diidentifikasi, sehingga mempermudah pengguna dalam mencari dan menemukan buku yang mereka butuhkan[9].

Penelitian ini akan menggunakan bahasa pemrograman Python untuk mengimplementasikan algoritma K-Means Clustering. Python dipilih karena berbagai pustaka dan alat bantu yang tersedia seperti Scikit-Learn, Pandas, dan Matplotlib, yang dapat mendukung proses analisis data, visualisasi, dan pengelompokan dengan efisien. Langkah-langkah implementasi meliputi pengumpulan data peminjaman buku yang mencakup informasi seperti judul buku, tanggal peminjaman, dan data demografis peminjam. Selanjutnya, data tersebut diolah dan disiapkan untuk proses clustering. Proses ini melibatkan normalisasi data, pemilihan jumlah cluster yang optimal, dan pelaksanaan algoritma K-Means Clustering. Hasil dari proses ini kemudian dianalisis untuk mendapatkan wawasan mengenai pola peminjaman. Evaluasi hasil clustering akan dilakukan menggunakan metode seperti Davies-Bouldin Index (DBI) untuk memastikan bahwa cluster yang terbentuk benar-benar representatif[10].

Dengan demikian, penerapan k-means clustering dalam analisis pola peminjaman buku perpustakaan tidak hanya memberikan manfaat langsung dalam pengelolaan perpustakaan, tetapi juga mendukung penelitian lebih lanjut dalam bidang data mining dan ilmu perpustakaan. Penelitian ini diharapkan dapat menjadi landasan bagi studi-studi berikutnya yang ingin mengembangkan metode analisis data yang lebih canggih dan aplikatif di berbagai bidang lain.

## **B. Metode Penelitian**

Dalam penelitian ini, dataset dianalisis melalui pre-processing yang mencakup pembuatan fitur baru, normalisasi, dan standard scaling untuk memastikan skala yang konsisten. Kemudian, algoritma K-Means diterapkan untuk mengelompokkan data ke dalam cluster berdasarkan jarak Euclidean. Evaluasi hasil dilakukan menggunakan Davies-Bouldin Index (DBI), yang menilai kualitas clustering dengan nilai lebih rendah menunjukkan hasil yang lebih baik. Adapun langkah-langkah ini dipaparkan secara visual pada gambar 1 berikut.



**Gambar 1.** Alur Penelitian

## 1. Dataset

Untuk menganalisis pola peminjaman buku menggunakan k-means clustering dengan 3 cluster, langkah awal adalah mengumpulkan dataset. Dataset yang digunakan bersumber dari platform seperti Kaggle dan mencakup informasi penting seperti tahun terbit, tanggal peminjaman, tanggal pengembalian, dan jumlah halaman. Dataset ini terdiri dari 1.323 data dan 12 kolom.

## 2. Pre-processing

Dalam tahap pre-processing, dilakukan dua langkah utama: pembuatan fitur dan normalisasi dengan standard scaling. Pembuatan fitur melibatkan pengembangan variabel baru dari data yang ada untuk memperkaya analisis, seperti menghitung durasi peminjaman dari tanggal peminjaman dan pengembalian. Sementara itu, normalisasi menggunakan standard scaling bertujuan untuk memastikan bahwa semua fitur memiliki mean nol dan deviasi standar satu, sehingga mencegah dominasi fitur dengan skala besar dalam analisis clustering.

## 3. Data Mining

Data mining adalah sebuah proses kompleks yang menggabungkan berbagai teknik dari pembelajaran mesin, kecerdasan buatan, dan disiplin ilmu lainnya untuk menganalisis data dalam jumlah besar dan mengekstraksi pengetahuan yang berguna[11]. Proses ini melibatkan identifikasi karakteristik penting dalam data, menemukan model yang menggambarkan kelas atau konsep, serta mengungkap pola dan hubungan antar atribut yang sebelumnya tidak terlihat. Data mining bukan hanya tentang mengolah data, tetapi juga tentang menghasilkan pengetahuan baru yang memiliki nilai tambah yang signifikan[12].

## 4. Clustering

Clustering, yang juga dikenal sebagai segmentasi, merupakan bagian penting dari data mining. Proses ini bertujuan untuk mengelompokkan sampel-sampel data

yang memiliki karakteristik serupa ke dalam kelompok-kelompok yang disebut cluster[13]. Setiap cluster terdiri dari sampel-sampel yang sangat mirip satu sama lain berdasarkan kriteria tertentu yang telah ditetapkan, tetapi sangat berbeda dari sampel-sampel yang terdapat dalam cluster lain. Dengan demikian, clustering membagi data menjadi beberapa subset di mana setiap subset berisi elemen-elemen yang homogen secara internal namun heterogen dibandingkan dengan elemen-elemen dari subset lainnya[14].

## 5. K-Means

K-Means adalah salah satu metode clustering data non-hirarkis yang mengelompokkan data ke dalam satu atau lebih cluster atau kelompok. Metode K-Means menggunakan algoritma clustering dengan pendekatan partisi, di mana pengelompokan didasarkan pada titik pusat atau centroid. Algoritma ini berbeda dari algoritma k-Medoids yang juga menggunakan pendekatan partisi tetapi didasarkan pada objek. Dalam K-Means, setiap cluster diwakili oleh centroid yang merupakan titik rata-rata dari semua data dalam cluster tersebut, dan proses ini terus berulang hingga tercapai kondisi optimal di mana data dalam cluster memiliki kesamaan yang tinggi[15].

Berikut adalah langkah-langkah algoritma K-Means, yaitu[11] :

1. Tentukan jumlah cluster (k) pada data set sebagai nilai centroid.
2. Menghitung jarak antara data dan titik pusat cluster menggunakan rumus-rumus dari Euclidian Distance. Yang dapat dilihat pada teori jarak Euclidian yang dirumuskan sebagai berikut:

$$D_e = \sqrt{(X_i - S_i)^2 + (Y_i - T_i)^2}$$

Dimana :

- $D_e$  adalah Euclidean Distance
  - $i$  adalah banyaknya objek
  - $(X, Y)$  adalah koordinat objek
  - $(S, T)$  adalah koordinat centroid
3. Pusat cluster baru akan ditentukan bila semua data telah ditetapkan dalam cluster terdekat.
  4. Proses penentuan titik pusat cluster dan penempatan data dalam cluster diulangi terus menerus sampai nilai centroid tidak berubah lagi.

## 6. Evaluasi Hasil

Dalam evaluasi hasil, Davies-Bouldin Index (DBI) digunakan untuk menilai kualitas clustering. DBI mengukur seberapa baik cluster terpisah dan seberapa kompak data di dalam setiap cluster. Nilai DBI yang lebih rendah menunjukkan bahwa cluster lebih terpisah dengan baik dan memiliki kedekatan yang lebih tinggi di dalam cluster, menandakan hasil clustering yang efektif.

## C. Hasil dan Pembahasan

### 1. Dataset

Dataset peminjaman buku dari perpustakaan, yang berisi sejumlah 1323 catatan peminjaman. Dataset ini terdiri dari 12 kolom yang mencakup informasi penting seperti tanggal peminjaman, tanggal pengembalian, jumlah halaman, dan tahun terbit buku. Gambar 2 merangkum deskripsi dari kolom-kolom yang ada dalam dataset.

| Book_Code | Book_Title                                | Author                | Year_Published | Pages | ID_Peminjaman | Tanggal_Peminjaman | Tanggal_Pengembalian | ID_Anggota Peminjam | Nama_Anggota |
|-----------|---|-----------------------|----------------|-------|---------------|--------------------|----------------------|---------------------|--------------|
| 0         | 911 The Confessions                       | Jean-Jacques Rousseau | 1781           | 606   | PB-2023-001   | 2024-09-05         | 2023-09-14           | 123894.0            | Felicity     |
| 1         | 882 The Art of Loving                     | Erich Fromm           | 1956           | 146   | PB-2023-002   | 2024-09-05         | 2023-09-14           | 123881.0            | Melania      |
| 2         | 102 The Adventures of Huckleberry Finn    | Mark Twain            | 1884           | 288   | PB-2023-003   | 2024-09-05         | 2023-09-14           | 123878.0            | Magdalena    |
| 3         | 378 My Life and Hard Times                | James Thurber         | 1933           | 115   | PB-2023-004   | 2024-09-05         | 2023-09-14           | 123879.0            | Anne-Marie   |
| 4         | 842 Romeo and Juliet                      | William Shakespeare   | 1596           | 146   | PB-2023-005   | 2024-09-05         | 2023-09-14           | 123880.0            | Bettie       |
| ...       | ...                                       | ...                   | ...            | ...   | ...           | ...                | ...                  | ...                 | ...          |
| 1318      | 1197 The Spiritual Growth of a Public Man | Abraham Lincoln       | 1973           | 47    | PB-2023-1320  | 2023-09-30         | 2023-11-01           | 125194.0            | Cynthia      |
| 1319      | 310 Saving Freedom                        | Joe Scarborough       | 2020           | 272   | PB-2023-1321  | 2023-09-30         | 2023-11-02           | 125189.0            | Martita      |
| 1320      | 742 NaN                                   | John Reed             | 1992           | 253   | NaN           | 2023-01-01 0:00:00 | 2022-12-31 0:00:00   | NaN                 | Adrian       |
| 1321      | 664 NaN                                   | Amos Oz               | 1885           | 272   | NaN           | 2023-01-01 0:00:00 | 2022-12-31 0:00:00   | NaN                 | Koressa      |
| 1322      | 250 NaN                                   | Horatio Alger, Jr.    | 1936           | 352   | NaN           | 2023-01-01 0:00:00 | 2022-12-31 0:00:00   | NaN                 | Joly         |

1323 rows x 12 columns

Gambar 2. Tampilan Dataset

## 2. Pre-processing

### 2.1 Feature Creation

Menambahkan kolom baru, yaitu Durasi\_Peminjaman, dilakukan dengan menghitung selisih antara kolom Tanggal\_Pengembalian dan Tanggal\_Peminjaman. Proses ini melibatkan konversi format tanggal menjadi tipe datetime yang sesuai, sehingga memungkinkan perhitungan durasi dalam satuan hari. Langkah ini merupakan bagian dari feature engineering yang bertujuan untuk menambah informasi yang relevan, meningkatkan akurasi analisis clustering, dan memberikan wawasan lebih mendalam mengenai pola peminjaman buku.

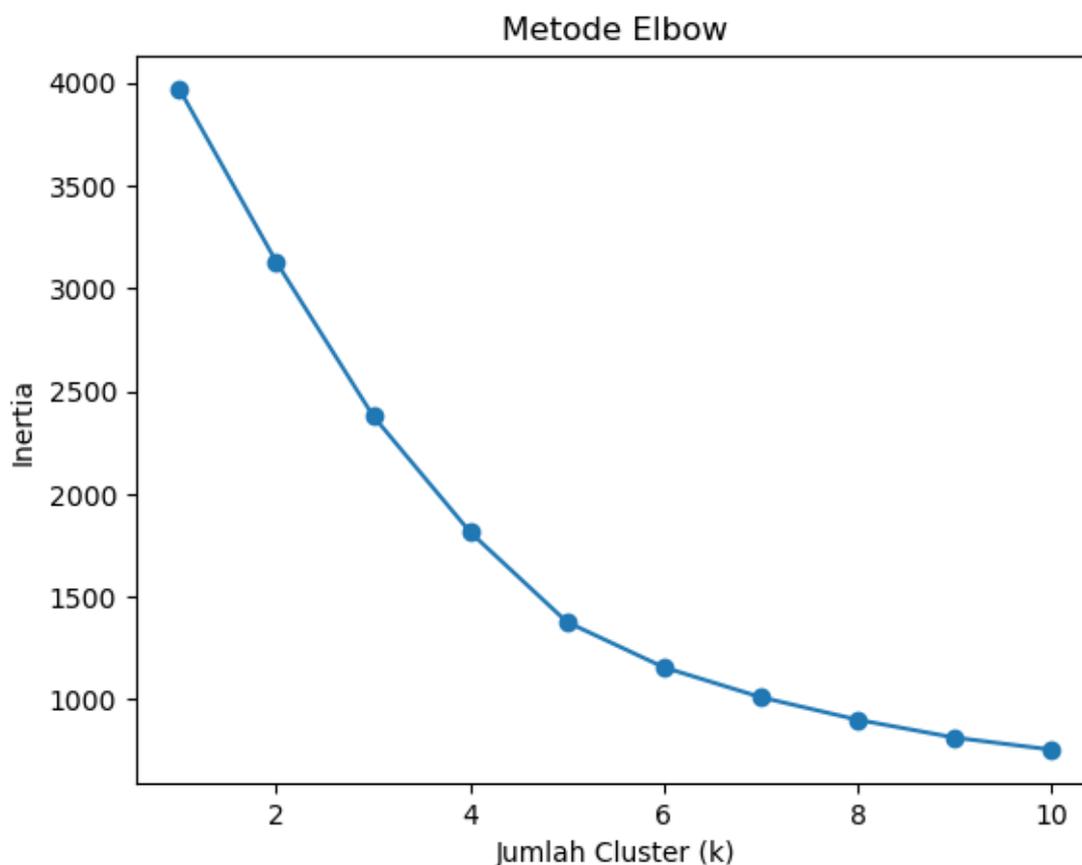
### 2.2 Normalisasi

Normalisasi menggunakan StandardScaler dari sklearn.preprocessing diterapkan untuk menormalkan fitur Durasi\_Peminjaman, Year\_Published, dan Pages dalam dataset. Proses normalisasi dilakukan dengan mengubah setiap fitur agar memiliki mean nol dan deviasi standar satu, memastikan skala yang konsisten di antara semua fitur. Hasil normalisasi ditambahkan kembali ke dataset sebagai kolom baru, memungkinkan analisis lebih lanjut dengan mempertimbangkan skala yang seragam. Langkah ini penting untuk mencegah dominasi fitur dengan skala yang lebih besar dalam analisis clustering.

## 3. Penerapan Metode dan Algoritma

### 3.1 Menentukan Jumlah Cluster

Untuk menentukan jumlah cluster yang optimal, digunakan Metode Elbow. Metode ini melibatkan plotting jumlah cluster versus Within-Cluster Sum of Squares (WCSS) atau inerti. Titik "elbow" dalam grafik adalah indikasi jumlah cluster yang optimal, di mana penurunan WCSS mulai melambat secara signifikan. Adapun hasil visualisasinya ada di gambar 3 berikut.



**Gambar 3.** Grafik Metode Elbow

Berdasarkan grafik diatas, titik elbow terlihat di sekitar  $k = 5$ , yang menunjukkan bahwa lima cluster mungkin merupakan jumlah optimal. Pada titik ini, penurunan inerti mulai melambat, menandakan keseimbangan antara kompleksitas model dan representasi data yang akurat. Oleh karena itu, memilih  $k = 5$  diharapkan memberikan pembagian cluster yang efektif dan informatif dalam analisis ini. Hasil clustering menunjukkan distribusi sebagai berikut:

Cluster 1: Buku yang diterbitkan lebih lama dengan jumlah halaman sedikit.

Cluster 2: Buku modern dengan peminjaman singkat.

Cluster 3: Buku klasik dengan durasi peminjaman lama.

Cluster 4: Buku dengan halaman banyak dan peminjaman moderat.

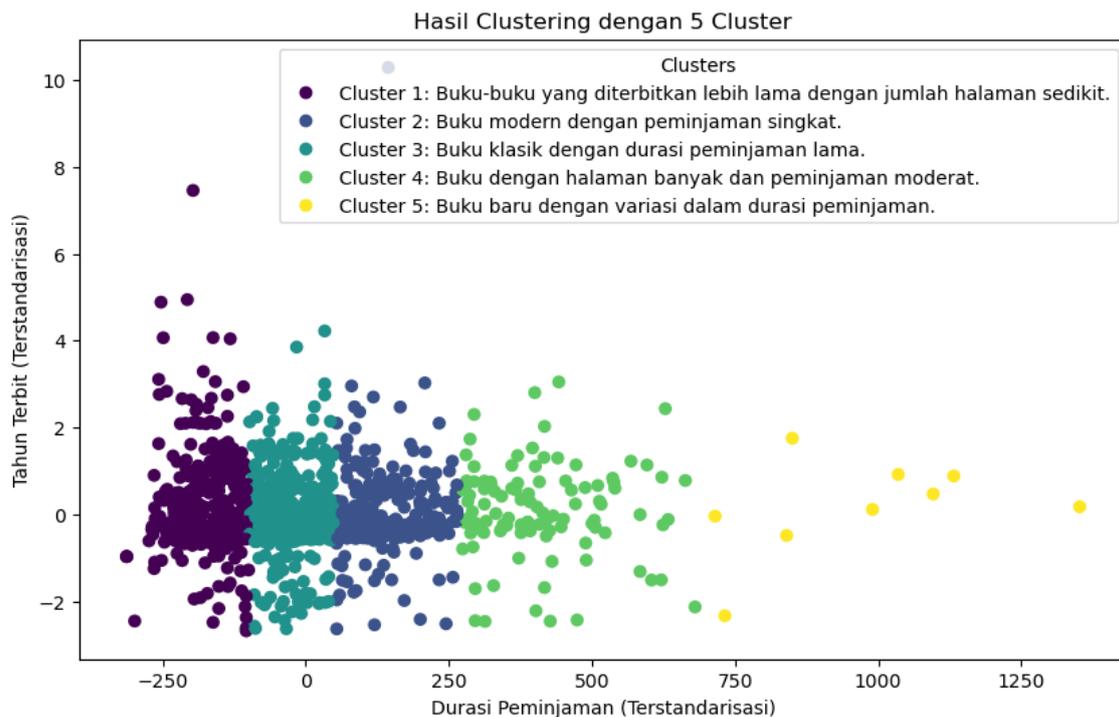
Cluster 5: Buku baru dengan variasi dalam durasi peminjaman.

### 3.2 Penerapan K-Means Clustering

Setelah menentukan jumlah cluster yang optimal, algoritma K-Means Clustering diterapkan pada data untuk membentuk 5 cluster. Distribusi cluster yang dihasilkan seperti tabel 1 berikut:

| Cluster   | Jumlah Anggota |
|-----------|----------------|
| Cluster 1 | 427            |
| Cluster 2 | 267            |
| Cluster 3 | 515            |
| Cluster 4 | 105            |
| Cluster 5 | 9              |

Untuk memberikan gambaran yang lebih jelas mengenai distribusi data dalam setiap cluster, berikut adalah scatter plot yang menunjukkan distribusi cluster berdasarkan tahun terbit buku, jumlah halaman dan durasi peminjaman.



**Gambar 4.** Hasil K-Means Clustering

**Cluster1:** Buku-buku yang diterbitkan lebih lama dengan jumlah halaman sedikit, menunjukkan preferensi untuk buku klasik atau historis dengan panjang yang lebih pendek.

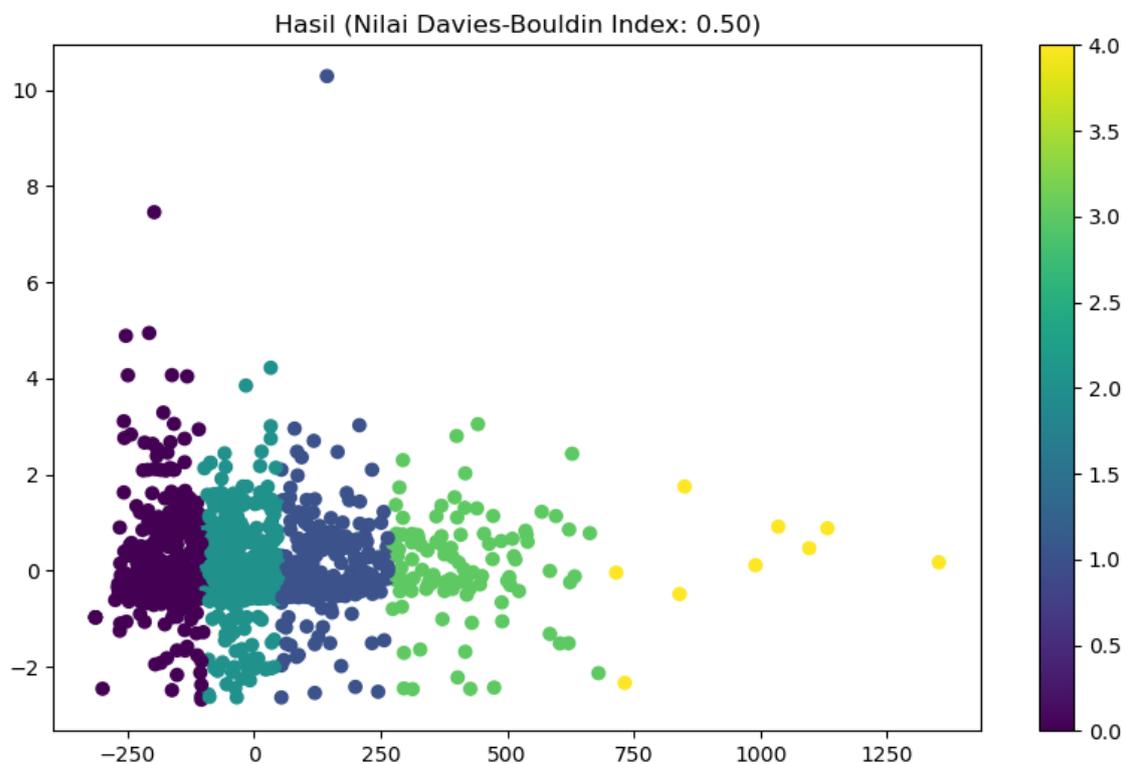
**Cluster2:** Buku modern dengan peminjaman singkat, mencerminkan tren bacaan terbaru yang cepat selesai dibaca.

**Cluster3:** Buku klasik dengan durasi peminjaman lama, menandakan karya klasik yang memerlukan waktu lebih lama untuk dibaca.

**Cluster4:** Buku dengan halaman banyak dan peminjaman moderat, menunjukkan bacaan yang lebih substansial tetapi dengan waktu pinjam yang tidak terlalu lama.

**Cluster5:** Buku baru dengan variasi dalam durasi peminjaman, mencerminkan variasi dalam preferensi dan waktu baca untuk buku-buku terbaru.

## 5. Evaluasi menggunakan Davies-Bouldin Index (DBI)



**Gambar 5.** Hasil Davies-Bouldin Index (DBI)

Dari gambar 5 di atas kita bisa melihat bahwa nilai DBI sebesar 0.50 mengindikasikan kualitas clustering yang baik, di mana cluster memiliki pemisahan yang jelas dan data di dalam cluster cukup kompak. DBI mengukur rasio jarak antar cluster terhadap jarak dalam cluster; semakin rendah nilai DBI, semakin baik definisi dan pemisahan cluster, mencerminkan efektivitas dari model clustering yang diterapkan. Sebelumnya, peneliti telah mencoba menggunakan tiga cluster dengan nilai DBI sebesar 1.16. Nilai DBI yang lebih tinggi tersebut menunjukkan bahwa cluster tidak terpisah dengan baik dan data di dalam cluster kurang kompak dibandingkan dengan hasil clustering yang menghasilkan nilai DBI 0.50. Temuan ini

menekankan bahwa penggunaan cluster yang lebih sedikit dengan nilai DBI yang lebih rendah dapat meningkatkan kualitas clustering secara signifikan.

#### D. Simpulan

Penelitian ini berhasil menganalisis pola peminjaman buku di perpustakaan menggunakan metode clustering dengan dataset yang terdiri dari 1.323 catatan peminjaman. Setelah melalui proses feature creation dan normalisasi, penentuan jumlah cluster optimal dilakukan menggunakan Metode Elbow yang menunjukkan bahwa lima cluster adalah yang paling efektif. Penerapan K-Means Clustering menghasilkan lima cluster dengan distribusi yang informatif, dan evaluasi menggunakan Davies-Bouldin Index (DBI) sebesar 0.50 mengindikasikan kualitas clustering yang baik. Hasil ini menegaskan bahwa penggunaan lima cluster dengan nilai DBI yang rendah dapat meningkatkan kualitas clustering secara signifikan, memberikan wawasan yang lebih mendalam mengenai pola peminjaman buku di perpustakaan.

#### E. Referensi

- [1] T. Silwattananusarn, P. Kulkanjanapiban, and K. L. Athakravisunthorn, "Mining and Analyzing Patron's Book-Loan Data and University Data to Understand Library Use Patterns," Dec. 2020.
- [2] N. N. Hasanah and A. S. Purnomo, "Implementasi Data Mining Untuk Pengelompokan Buku Menggunakan Algoritma K-Means Clustering (Studi Kasus: Perpustakaan Politeknik LPP Yogyakarta)," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 4, no. 2, pp. 300–311, Jul. 2022, doi: 10.47233/jteksis.v4i2.499.
- [3] F. Zhang, Y. Liu, C. Song, C. Yang, and S. Hong, "Empirical study of college students' extracurricular reading preference by functional data analysis of the library book borrowing behavior," *PLoS One*, vol. 19, no. 1 January, pp. 1–16, Jan. 2024, doi: 10.1371/journal.pone.0297357.
- [4] A. Sadiyah and H. Permana, "Manajemen Perpustakaan di Masa Pandemi Covid-19," *DIRASAH*, vol. 6, pp. 163–172, Feb. 2023.
- [5] Sasonoputri Fauzia Safitrie and Wahyusarib Retno, "Penerapan Algoritma Apriori Untuk Menemukan Pola Peminjaman Buku Di Perpustakaan," *SIMETRIS*, vol. 16, pp. 17–23, Jun. 2022.
- [6] I. A. Nur Afifah and H. Nurdiyanto, "Data Mining Clustering Dalam Pengelompokan Buku Perpustakaan Menggunakan Algoritma K-Means," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 8, no. 3, pp. 802–814, Aug. 2023, doi: 10.29100/jipi.v8i3.3891.
- [7] M. D. Chandra, E. Irawan, I. S. Saragih, A. P. Windarto, and D. Suhendro, "Penerapan Algoritma K-Means dalam Mengelompokkan Balita yang Mengalami Gizi Buruk Menurut Provinsi," *BIOS: Jurnal Teknologi Informasi dan Rekayasa Komputer*, vol. 2, no. 1, pp. 30–38, Mar. 2021, doi: 10.37148/bios.v2i1.19.
- [8] E. Bakker, "Implementasi Data Mining Clustering Data Perpustakaan Menggunakan Algoritma K-Means untuk Menentukan Penambahan Koleksi Buku di Perpustakaan UPY," 2020.

- [9] Royanti Nur Ika and Ismanto Bambang, "Pengelompokan Keaktifan Peminjaman Buku Di Perpustakaan Stmik Widya Pratama Dengan Metode K-Means," *IC - Tech*, vol. XV, pp. 81–84, Apr. 2020.
- [10] D. A. Fakhri, S. Defit, and Sumijan, "Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering," *Jurnal Informasi dan Teknologi*, vol. 3, pp. 160–166, Sep. 2021, doi: 10.37034/jidt.v3i3.137.
- [11] T. Tendean and W. Purba, "Analisis Cluster Provinsi Indonesia Berdasarkan Produksi Bahan Pangan Menggunakan Algoritma K-Means," *Jurnal Sains dan Teknologi*, vol. 1, no. 2, pp. 5–11, Mar. 2020.
- [12] J. Nasir, "Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means," *Jurnal SIMETRIS*, vol. 11, no. 2, 2020.
- [13] F. Amin, D. S. Anggraeni, and Q. Aini, "Penerapan Metode K-Means dalam Penjualan Produk Souq.Com," *Applied Information System and Management (AISM)*, vol. 5, no. 1, pp. 7–14, Apr. 2022, doi: 10.15408/aism.v5i1.22534.
- [14] L. Suriani, "Pengelompokan Data Kriminal Pada Poldasu Menentukan Pola Daerah Rawan Tindak Kriminal Menggunakan Data Mining Algoritma K-Means Clustering," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 1, no. 2, p. 151, Jan. 2020, doi: 10.30865/json.v1i2.1955.
- [15] W. Sudrajat, I. Cholid, and J. Petrus, "Penerapan Algoritma K-Means Clustering untuk Pengelompokan UMKM Menggunakan Rapidminer," Apr. 2022.