E-ISSN: 3031-8629
P-ISSN: 3032-3258
Halaman 4894-4911
Volume 3 Nomor 4 Tahun 2026

Clustering Peminjaman Buku Untuk Strategi Penambahan Koleksi: Komparasi K-*Means* Dan *Hierarchical Clustering*

[1)]Banu Dwi Putranto, [2)]Dzihni Safwa Alifah, [3)]Shalsa Sabrina Pungkasari

[1,2,3]Program Studi Teknologi Informasi, Universitas Amikom Purwokerto
e-mail: [1]banu@amikompurwokerto.ac.id

Abstrak

Perkembangan teknologi informasi telah mendorong transformasi dalam berbagai bidang, termasuk pengelolaan perpustakaan di perguruan tinggi. Perpustakaan tidak hanya menjadi tempat penyimpanan informasi, tetapi juga dituntut mampu menyediakan koleksi yang sesuai dengan kebutuhan pengguna. Salah satu tantangan yang dihadapi adalah menentukan koleksi buku mana yang perlu ditambahkan berdasarkan minat dan pola peminjaman mahasiswa. Penelitian ini bertujuan untuk memberikan rekomendasi penambahan koleksi buku di Perpustakaan Universitas Amikom Purwokerto dengan menerapkan metode *clustering* menggunakan algoritma K-*Means* dan *Hierarchical Clustering*. Data yang digunakan berupa transaksi peminjaman buku periode Januari–Desember 2025 sebanyak 1.224 record, yang disederhanakan menjadi 569 data setelah melalui tahap preprocessing. Atribut yang digunakan meliputi Judul Buku, Frekuensi Pinjam, dan Salinan. Penentuan jumlah cluster optimal dilakukan menggunakan metode *Elbow* dan *Silhouette*. Berdasarkan hasil perhitungan, nilai *Silhouette Score* tertinggi tercapai ketika jumlah *cluster* adalah 3, dengan skor sekitar 0.61. Hasil pengelompokan menunjukkan: *Cluster* 0 (14%) terdiri dari buku dengan frekuensi pinjam tinggi dan salinan sedang; *Cluster* 1 (57%) berisi buku dengan frekuensi pinjam rendah namun jumlah salinan tinggi; dan *Cluster* 2 (28%) terdiri atas buku dengan

frekuensi pinjam rendah dan salinan sedikit. Hasil penelitian ini menjadi dasar rekomendasi koleksi buku yang lebih relevan dan efisien bagi perpustakaan.

Kata kunci: Clustering, K-Means, Hierarchical, Koleksi Buku, Rekomendasi

Abstract

The development of information technology has driven transformation in various fields, including library management in universities. The library is not only a place to store information, but is also required to be able to provide collections that suit user needs. One of the challenges faced is determining which book collections need to be added based on student interests and borrowing patterns. This study aims to provide recommendations for adding book collections at the Amikom Purwokerto University Library by applying a clustering method using the K-Means and Hierarchical Clustering algorithms. The data used is in the form of book borrowing transactions for the period January–December 2025 totaling 1,224 records, which were simplified to 569 data after going through the preprocessing stage. The attributes used include Book Title, Borrowing Frequency, and Copies. Determination of the optimal number of clusters is carried out using the Elbow and Silhouette methods. Based on the calculation results, the highest Silhouette Score value is achieved when the number of clusters is 3, with a score of about 0.61. The grouping results show: Cluster 0 (14%) consists of books with high borrowing frequency and medium copies; Cluster 1 (57%) contains books with low borrowing frequency but a high number of copies; and Cluster 2 (28%) consists of books with low borrowing frequency and few copies. The results of this study serve as a basis for recommending book collections that are more relevant and efficient for libraries.

Keywords : *Clustering, K-Means, Hierarchical, Book Collection, Recommendation*

LATAR BELAKANG

Universitas Amikom Purwokerto memiliki perpustakaan sebagai pusat sumber belajar mahasiswa (Amanda Ayu Novitasari *et al.*, 2024). Di dalamnya tersedia beragam koleksi buku yang mendukung proses belajar mahasiswa, seperti buku panduan, skripsi, laporan tugas praktik, majalah, surat kabar, jurnal ilmiah, dan prosiding (Setyaningsih, 2019). Dari data peminjaman buku tersebut dapat digunakan untuk membantu pihak perpustakaan dalam mengelola koleksi buku yang tersedia di perpustakaan (Fatullah, 2025).

Pengelolaan sumber daya perpustakaan bertujuan untuk menyusun dan mengorganisir bahan pustaka agar dapat diakses dengan mudah serta ditempatkan sesuai sistem penataan tertentu, sehingga memudahkan pemustaka dalam mencari dan menggunakannya (Fatrizal, Arvianti and Sasono, 2025). Pengelolaan perpustakaan yang optimal akan memastikan ketersediaan koleksi yang relevan, fasilitas yang memadai, serta layanan yang berkualitas (Restia, Utami and Putri, 2024).

Penelitian ini mengidentifikasi permasalahan dalam pengelolaan koleksi buku yang dilakukan melalui wawancara langsung. Hasil wawancara menunjukkan bahwa keterbatasan ketersediaan buku menjadi salah satu kendala yang cukup sering terjadi, khususnya saat jumlah peminjaman lebih tinggi dari jumlah buku yang tersedia. Dari hasil wawancara dengan pustakawan perpustakaan Universitas Amikom Purwokerto, salah satu permasalahan yang sering muncul di perpustakaan adalah keterbatasan ketersediaan buku yang dipinjam, karena jumlah peminjaman tidak sebanding dengan jumlah buku yang ada. Menurut pustakawan, pernah terjadi ketidakseimbangan antara jumlah buku dan jumlah peminjam. Namun, apabila peminjam tidak menyampaikan secara langsung terkait kondisi buku yang dicari tersedia atau tidak, maka pustakawan tidak mengetahui bahwa buku tersebut jumlahnya kurang.

Dalam menyeimbangkan jumlah buku dan jumlah peminjam, penambahan jumlah koleksi buku terus dilakukan oleh pihak pengelola. Dalam proses rekomendasi penambahan koleksi buku selama ini hanya didasarkan pada pengajuan yang dilakukan mahasiswa melalui website student, tanpa mempertimbangkan data transaksi peminjaman yang lebih merepresentasikan minat terhadap buku-buku yang sering

digunakan. Menurut pustakawan, analisis peminatan mahasiswa berdasarkan data transaksi peminjaman belum dapat dikelola secara optimal karena prosesnya masih dilakukan secara manual satu per satu, sehingga memerlukan waktu yang cukup lama dan kurang efisien. Dalam mengatasi permasalahan tersebut, perlu dilakukan sebuah metode analisis yang mampu mengelola data dalam jumlah besar salah satunya dengan menggunakan teknik data mining.

Teknik data mining mencakup berbagai metode seperti klasifikasi, regresi, asosiasi (Fauzi Prameswari and Latifah Dwi Mutiara Sari, 2024). Salah satunya yang relevan untuk pengelolaan koleksi perpustakaan adalah clustering (Rosiana and Yuadi, 2025). Pada metode clustering, K-Means dan Hierarchical Clustering merupakan algoritma yang sering digunakan dalam pengelompokkan data (Alif *et al.*, 2024). Popularitas algoritma K-Means dalam dunia penelitian tidak terlepas dari kemampuannya mengelompokkan data secara efisien (Padiku and Lahinta, 2022). Kelebihan dari algoritma ini adalah efektivitasnya dalam menangani dataset berukuran besar serta kesesuaiannya untuk klaster yang memiliki bentuk globular. Algoritma K-Means memiliki beberapa keterbatasan seperti sensitivitas terhadap penentuan jumlah klaster (k) di awal proses. *Hierarchical clustering* adalah salah satu teknik dalam *machine learning* yang digunakan untuk mengelompokkan data ke dalam kelompok-kelompok yang lebih kecil berdasarkan kemiripan mereka (Musdalifah M. Ramly, Sudarmin and Poerwanto, 2022). Metode ini sering digunakan dalam analisis data eksploratori untuk menemukan struktur dalam dataset tanpa perlu supervise (Zuhal, 2022).

Dalam memperoleh hasil yang akurat maka perlu adanya perbandingan algoritma clustering yaitu menggunakan k-*means* dan *Hierarchical Clustering*. Penelitian sebelumnya menghasilkan klaster optimal, di mana *cluster*-1 mencakup dua jenis klasifikasi buku yang paling sering dipinjam, sedangkan cluster-2 terdiri dari 51 jenis klasifikasi buku yang jarang atau kurang diminati untuk dipinjam (Mau, Chrisinta and Binsasi, 2024). Penelitian ini bertujuan untuk menerapkan algoritma *clustering* K-*Means* dan *Hierarchical Clustering* dalam rekomendasi penambahan koleksi buku di

perpustakaan Universitas Amikom Purwokerto, dengan fokus pada analisis pola peminjaman buku untuk menentukan koleksi buku yang perlu ditambah sesuai dengan kebutuhan pengguna perpustakaan.

METODE

Metode penelitian berisi tahapan urutan penelitian sebagai berikut.

2.1 Pengumpulan Data

Penelitian ini menggunakan data koleksi buku dan data transaksi peminjaman buku yang diperoleh dari Perpustakaan Universitas Amikom Purwokerto. Pengumpulan data dilakukan melalui wawancara langsung dengan pustakawan. Melalui proses wawancara tersebut, diperoleh data koleksi buku dan data transaksi peminjaman buku selama periode 1 tahun, dengan total sebanyak 1224 *record*. Rentang waktu pengambilan data dimulai dari Januari 2025 hingga Desember 2025.

2.2 *Preprocessing* Data

Tahap ini dilakukan preprocessing data yang dimulai dengan penentuan dataset awal, di mana dipilih atribut-atribut yang relevan dari data koleksi buku dan data transaksi peminjaman buku. Selanjutnya dilakukan tahap transformasi data, yang mencakup penggabungan data dengan judul buku yang sama namun jumlah peminjam yang berbeda, serta pembersihan data dari karakter seperti tanda koma (,) dan tanda petik (') untuk memastikan data dapat diproses oleh perangkat lunak Weka. Tahap akhir dari *preprocessing* adalah menentukan dataset akhir. Setelah proses penentuan dataset awal dan transformasi selesai, data kemudian diubah dari format CSV atau Excel menjadi format .arff agar siap digunakan dalam proses analisis penelitian.

2.3 Penggunaan *K-Means* dan *Hierarchy Clustering*

Penelitian ini akan menerapkan metode *cluster*ing menggunakan algoritma *K-Means* dan Hierarchy *Clustering* untuk mengelompokkan koleksi buku berdasarkan tingkat peminatan mahasiswa, dengan tujuan memberikan rekomendasi dalam pengelolaan buku di Perpustakaan Amikom Purwokerto. Proses ini dilakukan dengan pendekatan perhitungan jarak menggunakan rumus *Euclidean Distance*, yang

memungkinkan sistem untuk menentukan kedekatan antar data (buku) berdasarkan pola peminjaman atau ketertarikan pengguna.

Tahap awal dalam metode ini adalah menentukan jumlah kelompok (*cluster*) yang akan dibentuk. Untuk memperoleh jumlah *cluster* yang paling optimal, digunakan dua pendekatan umum, yaitu metode *Elbow* dan metode *Silhouette*. Metode *Elbow* dilakukan dengan mengamati perubahan nilai varians dalam masing-masing *cluster* seiring bertambahnya jumlah *cluster*, titik siku (elbow) pada grafik menunjukkan jumlah *cluster* yang paling efisien. Sementara itu, metode *Silhouette* menilai seberapa baik suatu data berada dalam *cluster*-nya sendiri dibandingkan dengan *cluster* lain. Nilai *Silhouette* yang tinggi menunjukkan bahwa konfigurasi *cluster* tersebut merupakan yang paling ideal.

Tahap selanjutnya sistem akan memilih pusat *cluster* awal secara acak. Kemudian dilakukan perhitungan jarak setiap buku terhadap pusat *cluster* tersebut. Buku-buku dikelompokkan berdasarkan jarak terdekat ke pusatnya, lalu dilakukan pembaruan posisi pusat *cluster* berdasarkan rata-rata posisi data dalam masing-masing kelompok. Proses ini akan terus diulang hingga pusat *cluster* tidak mengalami perubahan lagi. Hasil pengelompokan ini divisualisasikan dalam bentuk grafik *scatter plot*, yang menampilkan persebaran data berdasarkan frekuensi peminjaman dan jumlah salinan untuk masing-masing *cluster*, sehingga memudahkan analisis visual terhadap pola peminjaman buku.

2.4 Penarikan Kesimpulan dan Rekomendasi.

Kesimpulan dari proses perhitungan WEKA didapatkan hasil analisis yang diperoleh dari penerapan algoritma *K-Means* dan *hierarchy*, yang menghasilkan pengelompokan terbaik berdasarkan data transaksi peminjaman buku. Hasil pengelompokan ini digunakan sebagai dasar untuk menentukan jenis-jenis buku yang direkomendasikan, sehingga dapat dijadikan acuan dalam penambahan koleksi buku di Perpustakaan Universitas Amikom Purwokerto. Tahap ini menetapkan ketentuan yang digunakan sebagai dasar dalam memberikan rekomendasi penambahan koleksi buku berdasarkan hasil analisis

HASIL DAN PEMBAHASAN

3.1 Preprocessing Data

3.1.1 Dataset Awal

Tahapan ini mencakup proses pengurangan atribut yang dianggap tidak relevan agar data yang dianalisis menjadi lebih terfokus dan ringkas. Berdasarkan dataset awal terdiri atas beberapa atribut seperti biblio_id, judul, salin, tempat terbit, penerbit, ISBN/ISSN, id anggota, nama anggota, dll. Namun, dalam proses analisis menggunakan algoritma *K-Means* dan *Hierarchy*, hanya atribut bertipe numerik yang relevan dan akan digunakan. Atribut yang dipilih untuk dianalisis melalui WEKA adalah judul buku, frekuensi pinjam, dan salinan. Berikut merupakan hasil dari proses seleksi atribut.

Tabel 1. Hasil Pemangkasan Atribut

Judul_Buku	**Frekuensi_Pinjam**	**Salinan**
#About Life	1	3
#Nasihat Diri Untuk Para Pekerja	1	2
#Pernah Tenggelam	1	3
…	…	…
…	…	…
Yesterday In Bandung	1	2
You Do You; Discovering Life Through Experiments & Self-Awareness	1	4
Young Sherlock Holmes Red Leech	1	1

3.1.2 Transformasi Data

Tahapan ini mengintegrasikan data-data tersebut ke dalam satu entri. Selain proses penggabungan, dilakukan pula penyesuaian pada penulisan nama item yang mengandung tanda baca seperti koma (,), (#), (&) dan tanda petik satu (’). Penyesuaian ini bertujuan agar data dapat dikenali dan diproses dengan baik oleh perangkat lunak Weka.

Tabel 2. Hasil Sesudah Pengubahan Nama & Penggabungan Data

Judul_Buku	Frekuensi_Pinjam	Salinan
About Life	1	3
Nasihat Diri Untuk Para Pekerja	1	2
Pernah Tenggelam	1	3
...	...	...
...	...	...
Yesterday In Bandung	2	2
You Do You; Discovering Life Through Experiments And Self-Awareness	10	4
Young Sherlock Holmes Red Leech	1	1

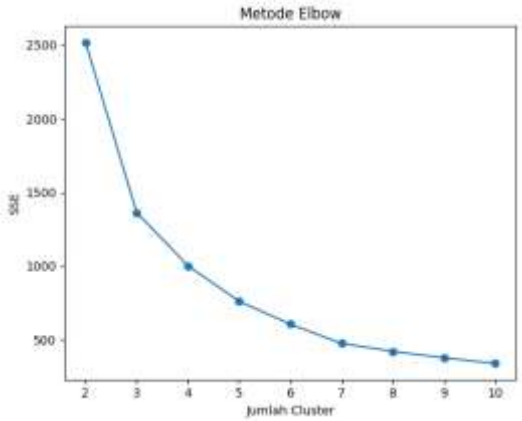
3.1.3 Dataset Akhir

Hasil seleksi pada dataset awal menetapkan tiga atribut utama yang digunakan, yaitu Judul Buku, Frekuensi_Pinjam dan Salinan. Selama proses transformasi data, jumlah entri mengalami pengurangan signifikan, dari semula 1.224 item menjadi 569 item. Hal ini terjadi akibat proses konsolidasi terhadap item yang memiliki judul serupa namun mencatat angka peminjaman yang berbeda. Pada tahap ini juga dilakukan konversi format data dari file berekstensi CSV ke dalam format .arff agar data dapat digunakan dalam proses analisis menggunakan perangkat lunak Weka.

3.2 Penggunaan K-Means dan Hierarchy Clustering

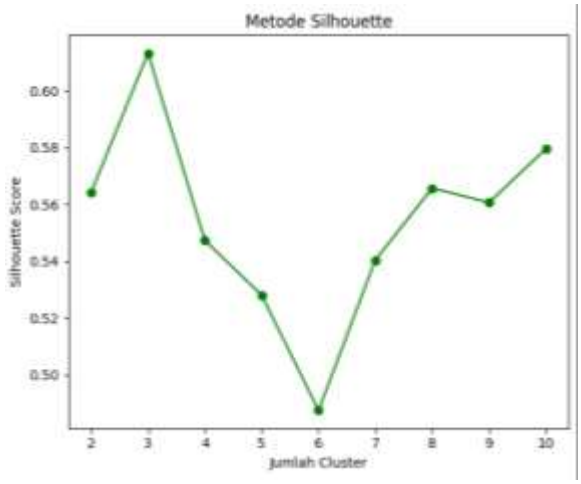
3.2.1 Menentukan Jumlah *Cluster*

Proses *clustering* dengan algoritma *K-Means* dan *Hierarchy* diawali dengan penentuan jumlah *cluster*. Tahap ini dilakukan setelah data siap dianalisis melalui serangkaian tahap *preprocessing* dan penetapan atribut numerik. Untuk mendapatkan hasil *cluster*ing yang optimal, terlebih dahulu dilakukan analisis penentuan jumlah *cluster* terbaik menggunakan metode *Elbow* dan *Silhouette* melalui platform Google Colaboratory (Colab).



Gambar 1. Grafik Metode Elbow

Hasil visualisasi menggunakan metode Elbow pada gambar 1 menunjukkan bahwa grafik mengalami penurunan tajam dari jumlah *cluster* 2 ke 3, lalu ke 4. Setelah titik tersebut, penurunan nilai SSE (*Sum of Squared Errors*) mulai melandai. Titik perubahan arah yang signifikan atau yang disebut sebagai "titik siku" (elbow) terlihat paling mencolok pada jumlah *cluster* ke-3. Hal ini mengindikasikan bahwa jumlah *cluster* optimal adalah sebanyak tiga kelompok, karena setelahnya tidak terjadi lagi penurunan nilai SSE yang signifikan.



Gambar 2. Grafik Metode Silhoutte

Metode Silhouette digunakan untuk mengevaluasi kualitas pemisahan antar *cluster*. Berdasarkan hasil pada gambar 2 nilai Silhouette *Score* tertinggi tercapai ketika

jumlah *cluster* adalah 3, dengan skor sekitar 0.61. Skor ini menunjukkan bahwa struktur pengelompokan paling stabil dan paling terpisah terjadi saat data dibagi menjadi tiga *cluster*. Setelah jumlah *cluster* melebihi 3, nilai Silhouette *Score* cenderung menurun atau tidak menunjukkan peningkatan yang signifikan, yang menandakan bahwa kualitas *cluster* menurun jika dipaksakan untuk dibagi lebih banyak. Berdasarkan analisis dari kedua metode, baik Elbow maupun Silhouette, disimpulkan bahwa jumlah *cluster* optimal untuk data peminjaman buku ini adalah sebanyak 3 *cluster.*

3.2.2 Menentukan Pusat *Cluster* Awal

Tahap ini WEKA memilih tiga pusat *cluster* secara acak dari dataset yaitu:

Cluster 0: (Frekuensi_Pinjam = 3, Salinan = 3)

Cluster 1: (Frekuensi_Pinjam = 1, Salinan = 3)

Cluster 2: (Frekuensi_Pinjam = 1, Salinan = 2)

3.2.3 Menghitung Jarak Objek ke Pusat *Cluster*

Perhitungan jarak antara masing-masing objek dan titik pusat klaster dilakukan dengan menggunakan rumus *Euclidean Distance*, Tahap awal iterasi pertama, dilakukan perhitungan jarak antara data pertama dan pusat klaster pertama (*Cluster*0). Perhitungan tersebut dinyatakan sebagai berikut:

$$D1 = \sqrt{(1-3)^2 + (3-3)^2} = 2$$

Hasil dari perhitungan tersebut menunjukkan bahwa jarak antara data pertama dan pusat klaster pertama adalah 2. Tahapan selanjutnya menghitung jarak antara data pertama dan pusat klaster kedua (*Cluster*1), sebagai berikut:

$$D1 = \sqrt{(1-1)^2 + (3-3)^2} = 0$$

Hasil dari perhitungan tersebut menunjukkan bahwa jarak antara data pertama dan pusat klaster pertama adalah 0. Proses selanjutnya melibatkan penghitungan jarak antara data pertama dan pusat klaster ketiga (*Cluster2*) guna mengevaluasi kedekatan relatif antar data dan centroid, yang dihitung sebagai berikut:

$$D1 = \sqrt{(1-1)^2 + (3-2)^2} = 1$$

Hasil perhitungan tersebut, diketahui bahwa jarak antara data pertama dan pusat klaster ketiga adalah 1. Proses perhitungan ini dilakukan secara berurutan dari data ke-1 hingga data ke-569.

3.2.4 Mengelompokkan Data Berdasarkan Jarak Minimum

Hasil perhitungan jarak antar data dan pusat klaster digunakan oleh WEKA untuk menentukan keanggotaan klaster berdasarkan jarak terpendek. Berikut hasil pengelompokannya dapat dilihat pada tabel 3:

Tabel 3 Hasil Pengelompokkan Data Iterasi

Cluster	Jumlah Data	Persentase
0	82	14%
1	326	57%
2	161	28%

3.2.5 Menghitung Pusat *Cluster* Baru

Tahapan selanjutnya menghitung ulang posisi pusat *cluster* (centroid) dengan menggunakan rata-rata nilai dari masing-masing atribut numerik dalam setiap *cluster*. Berikut hasil perhitungan pusat *cluster* baru iterasi dapat dilihat pada tabel 4:

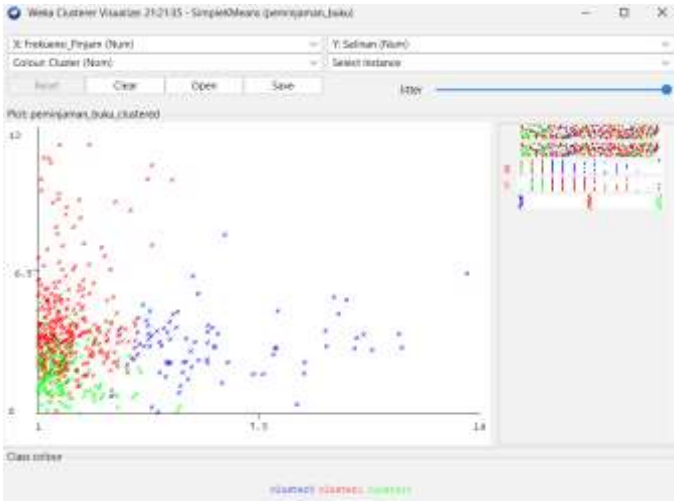
Tabel 4 Hasil Perhitungan Pusat *Cluster* Baru Iterasi

Atribut	*Cluster* 0	*Cluster* 1	*Cluster* 2
Frekuensi_Pinjam	6.0854	1.5429	1.3789
Salinan	3.1463	3.9724	1.7391

Hasil penghitungan pusat *cluster* baru pada iterasi yang ditampilkan oleh WEKA, diperoleh nilai pusat *cluster* baru untuk masing-masing kelompok. *Cluster* 0 memiliki rata-rata frekuensi peminjaman 6.08 kali dan jumlah salinan 3.14. *Cluster* 1 memiliki rata-rata frekuensi peminjaman 1.54 kali dan jumlah salinan 3.97. *Cluster* 2 memiliki rata-rata frekuensi peminjaman 1.37 kali dan jumlah salinan 1.73.

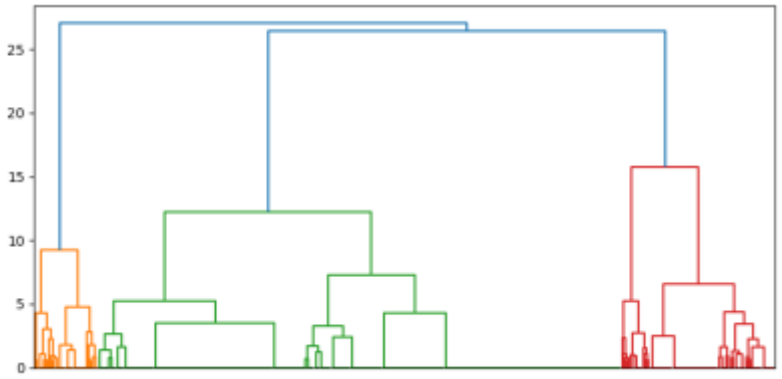
3.2.6 Visualisasi Data

Penelitian ini berhasil melakukan analisis dengan jumlah data yang dianalisis sebanyak 569 data. Hasil dari proses pengelompokan tersebut ditampilkan pada grafik *cluster*ing pada gambar 5:



Gambar 5 Hasil Grafik k-maens *Clustering*

Gambar 5 merupakan hasil dari proses perhitungan menggunakan weka, diketahui bahwa pada *cluster* 0 terdapat 82 item dengan prosentase sebesar 14%, pada *cluster* 1 terdapat 326 item dengan prosentase sebesar 57% dan pada *cluster* 2 terdapat 161 item yang memiliki prosentase sebesar 28%. Selanjutnya adalah hasil grafik menggunakan hierarchy clustering pada gambar 6.



Gambar 6 Hasil Grafik *Hierarchy Clustering*

Berdasarkan dendrogram tersebut, data koleksi buku dapat dibagi menjadi **tiga cluster utama**. Pemotongan dendrogram pada ketinggian sekitar **15–18** menghasilkan tiga kelompok buku yang memiliki karakteristik berbeda berdasarkan **jumlah salinan** dan **jumlah peminjam**.

3.2.7 Pengelompokan Judul Buku

Berikut ini adalah hasil pengelompokan judul buku pada tabel 4.

Tabel 4 Hasil *Cluster*ing *K-Means* pada Weka

No	Judul_Buku	Frekuensi_ Pinjam	Salinan	*Cluster*
1	About Life	1	3	*cluster*1
2	Nasihat Diri Untuk Para Pekerja	1	2	*cluster*2
3	Pernah Tenggelam	1	3	*cluster*1
4	1 Juta Rupiah Pertama Anda Dari Affiliate Marketing; Panduan Dan Langkah Praktis Menghasilkan Rupiah Dari Internet	2	3	*cluster*1
5	1000 Mindset Berpikir Positif	9	1	*cluster*0
…	…	…	…	…
…	…	…	…	…
…	…	…	…	…
565	Why? Advanced Material Material Maju	2	3	*cluster*1
566	Wingit	5	2	*cluster*0

567	Yesterday In Bandung	2	2	*cluster*2
568	You Do You; Discovering Life Through Experiments And Self-Awareness	10	4	*cluster*0
569	Young Sherlock Holmes Red Leech	1	1	*cluster*2

Proses *clustering* dengan algoritma *K-Means* menghasilkan tiga klaster utama yang merepresentasikan pengelompokan koleksi buku berdasarkan frekuensi peminjaman dan jumlah salinan. Hasil *clustering* digunakan untuk mengidentifikasi buku-buku yang layak direkomendasikan untuk penambahan. Berikut adalah tabel hasil rekomendasi:

Tabel 5 Rekomendasi Penambahan Buku berdasarkan k-*maens clustering*

No	Judul Buku	Rekomendasi
1	1000 Mindset Berpikir Positif	Prioritas Tinggi
2	Manajemen Proyek	Sangat Direkomendasi
3	Metode Penelitian Kuantitatif	Sangat Direkomendasi
4	Bumi	Sangat Direkomendasi
5	Sesuk	Sangat Direkomendasi
6	Metode Penelitian Kuantitatif Kualitatif dan R&D	Sangat Direkomendasi
7	Pengenalan Machine Learning dengan Python	Direkomendasi
8	Sejarah Dunia yang Disembunyikan	Direkomendasi
9	You Do You	Direkomendasi
10	Manajemen Proyek Berbasis Teknologi Informasi	Direkomendasi
11	Mindset	Direkomendasi

12	Duduk Dulu Jangan Lupa Jadi Manusia	Direkomendasi

Sedangkan hasil rekomendasi dari analisis *hierarchy* clustering sebagai berikut.

Tabel 6 Rekomendasi Penambahan Koleksi Buku berdasarkan *hierarchy clustering*

No	Judul Buku	Rekomendasi
1	1000 Mindset Berpikir Positif	Prioritas Tinggi
2	Manajemen Proyek	Sangat Direkomendasi
3	Metode Penelitian Kuantitatif	Sangat Direkomendasi
4	Sesuk	Sangat Direkomendasi
5	Bumi	Sangat Direkomendasi
6	Metode penelitian kuantitatif kualitatif dan R&D	Sangat Direkomendasi
7	Pengenalan Machine Learning dengan Python	Direkomendasi
8	Sejarah Dunia yang Disembunyikan	Direkomendasi
9	You Do You; Discovering Life Through Experiment...	Direkomendasi
10	Manajemen Proyek Berbasis Teknologi Informasi	Direkomendasi
11	Duduk Dulu Jangan Lupa Jadi Manusia	Direkomendasi
12	Mindset	Direkomendasi

SIMPULAN

Hasil pemrosesan yang telah dilakukan menggunakan algoritma *K-Means Clustering* pada aplikasi WEKA, sebanyak 569 data berhasil dikelompokkan ke dalam 3 cluster dengan total 5 iterasi. Berdasarkan hasil pengelompokan, terdapat tiga klaster

buku dengan karakteristik berbeda, *cluster* 0 yang mencakup 14% dari total, terdiri atas buku-buku dengan tingkat peminjaman yang tinggi (rata-rata 6,08) dan jumlah salinan yang tergolong sedang (rata-rata 3,14 salinan). Sementara itu, *cluster* 1 mencakup sebagian besar koleksi (57%) dengan karakteristik frekuensi peminjaman rendah (rata-rata 1,54) namun memiliki jumlah salinan yang relatif banyak (rata-rata 3,97). Adapun *cluster* 2, yang mencakup 28% dari total buku, terdiri dari koleksi dengan frekuensi peminjaman rendah (rata-rata 1,37) serta jumlah salinan yang juga sedikit (rata-rata 1,73). Hasil perbandingan algoritma k-*means* dan *hierarchy clastering* berdasarkan rekomendasi prioritas dan sangat direkomendasikan hasil judul bukunya sama. Peneliti selanjutnya diharapkan dapat mencoba atau menggabungkan metode clustering lain, seperti DBSCAN untuk melihat perbedaan dan juga perbandingan hasil pengelompokan dan meningkatkan keakuratan dalam sistem rekomendasi yang dihasilkan.

UCAPAN TERIMA KASIH

ucapan terimakasih peneliti tunjukan kepada LPPM Universitas Amikom Purwokerto yang telah mendanai penelitian ini melalui program Penelitian Dosen Muda Amikom (PDMA) tahun 2026.

DAFTAR PUSTAKA

Alif, A. *et al.* (2024) "Penerapan Algoritma K-means Clustering dan Hierarchical Clustering dalam Mengelompokkan Data Pengangguran di Karawang," *Jurnal Algoritma*, 21(2), pp. 209–220. Available at: https://doi.org/10.33364/algoritma/v.21-2.2155.

Amanda Ayu Novitasari et al. (2024) "Analisis Manajemen Layanan ELLS Universitas Amikom Purwokerto Menggunakan Framework ITIL V3," Switch : Jurnal Sains dan Teknologi Informasi, 2(6), pp. 24–35. Available at: https://doi.org/10.62951/switch.v2i6.278.

Fatrizal, Arvianti, E.Y. and Sasono, A.D. (2025) "Optimalisasi Pengelolaan Perpustakaan sebagai UpayaPeningkatan Pelayanan Publik Provinsi Nusa Tenggara Barat,"

LENVARI: Jurnal of Social Science, 3(2), pp. 53–66. Available at: https://doi.org/.

Fatullah, A.G. (2025) Pengelompokan Buku Berdasarkan Tingkat Keterpakaian Menggunakan Algoritma K-means Clustering Pada Perpustakaan MAN 2 Kota Tangerang. Universitas Islam Negeri Maulana Malik Ibrahim. Available at: http://etheses.uin-malang.ac.id/78069/.

Fauzi Prameswari, S. and Latifah Dwi Mutiara Sari, N. (2024) "Literatur Review: Implementasi Metode dan Dataset pada Data Mining," Seminar Nasional Informatika-FTI UPGRIS, 2, pp. 531–539.

Mau, P.Y., Chrisinta, D. and Binsasi, E. (2024) "Implementasi Data Mining Dalam Menentukan Penambahan Koleksi Buku di Perpustakaan: Algoritma K-Means Clustering," Journal of Mathematics, Computations and Statistics, 7(1), pp. 122–132. Available at: https://doi.org/10.35580/jmathcos.v7i1.1938.

Musdalifah M. Ramly, Sudarmin, S. and Poerwanto, B. (2022) "Analisis Hierarchical Clustering Multiscale Bootstrap (Kasus: Indikator Kemiskinan di Provinsi Sulawesi Selatan Tahun 2020)," VARIANSI: Journal of Statistics and Its application on Teaching and Research, 4(3), pp. 142–152. Available at: https://doi.org/10.35580/variansiunm26.

Padiku, I. and Lahinta, A. (2022) "Penerapan Clustering K-Means Untuk Mendukung Pengelolaan Koleksi Pada Perpustakaan Fakultas Teknik Universitas Negeri Gorontalo," Jurnal Teknik, 20(1), pp. 54–62. Available at: https://doi.org/10.37031/jt.v20i1.206.

Restia, Utami, R.N. and Putri, A.D. (2024) "Manajemen Pengolahan Bahan Pustaka Di Perpustakaan SMA Negeri 3 Palembang," IQRA': Jurnal Ilmu Perpustakaan Dan Informasi, 2(1), pp. 306–312.

Rosiana, L. and Yuadi, I. (2025) "K-Means Clustering untuk Analisis Tren Peminjaman Buku di Perpustakaan," Journal of Technology and Informatics (JoTI), 7(1), pp. 1–10. Available at: https://doi.org/10.37802/joti.v7i1.933.

Setyaningsih, G. (2019) "Pengaruh Persepsi Kebermanfaatan dan Persepsi Kemudahan Penggunaan Terhadap Penggunaan Sistem Secara Aktual pada Sistem Informasi Perpustakaan STMIK Amikom Purwokerto," Citisee.Amikompurwokerto.Ac.Id. Purwokerto, pp. 248–253. Available at: https://citisee.amikompurwokerto.ac.id/assets/proceedings/2017/SI14.pdf.

Zuhal, N.K. (2022) "Study Comparison K-Means Clustering dengan Algoritma Hierarchical Clustering," Prosiding Seminar Nasional Teknologi Dan Sains, pp. 200–205.