



Mapping Library User Behavior Base On K-Means Clustering Of Ma'soem University Student

Pemetaan Perilaku Pengguna Perpustakaan Berbasis K-Means Pada Mahasiswa Universitas Ma'soem

Sofia Dewi¹, Ai Kresnawati², Sinta Sopyanti³, Alma Sulmainah⁴

^{1,2,3,4} Sistem Informasi, Universitas Ma'soem, Indonesia

E-Mail: ¹sofiadewi@masoemuniversity.ac.id, ²kresnawatia@gmail.com, ³sinta.sopyanti54321@gmail.com,
⁴almasulmainah26@gmail.com

*Makalah: Diterima 07 Agustus 2025 ; Diperbaiki 24 Agustus 2025; Disetujui 25 Agustus 2025
Corresponding Author: Sofia Dewi*

Abstrak

Perpustakaan memiliki peran strategis dalam mendukung kegiatan akademik di perguruan tinggi, namun rendahnya tingkat kunjungan dan peminjaman buku oleh mahasiswa menunjukkan adanya kesenjangan perilaku pengguna. Penelitian ini bertujuan untuk memetakan pola perilaku mahasiswa dalam memanfaatkan layanan perpustakaan dengan menerapkan algoritma K-Means Clustering terhadap data kunjungan dan peminjaman selama tiga bulan terakhir. Melalui pendekatan data mining, mahasiswa dikelompokkan ke dalam tiga kluster: aktif, cukup aktif, dan pasif, berdasarkan frekuensi kunjungan serta jumlah peminjaman buku. Proses analisis dilakukan menggunakan perangkat lunak RapidMiner dengan tahapan meliputi data preprocessing dan evaluasi model menggunakan indeks Davies-Bouldin. Hasil penelitian menunjukkan bahwa pendekatan klusterisasi efektif dalam mengidentifikasi segmentasi perilaku mahasiswa, yang selanjutnya dapat digunakan sebagai dasar dalam merumuskan kebijakan pengembangan layanan perpustakaan yang adaptif dan terarah. Penelitian ini memberikan kontribusi berupa kerangka kerja berbasis data yang dapat direplikasi oleh institusi pendidikan lain guna meningkatkan efektivitas dan efisiensi pengelolaan perpustakaan.

Kata Kunci : Data Mining, K-Means, Klusterisasi, Perilaku Mahasiswa, RapidMiner

Abstract

Libraries have a strategic role in supporting academic activities in higher education, but the low level of visits and book borrowing by students indicates a gap in user behavior. This study aims to map student behavior patterns in utilizing library services by applying the K-Means Clustering algorithm to visit and borrow data for the last three months. Through a data mining approach, students are grouped into three clusters: active, quite active, and passive, based on the frequency of visits and the number of book borrowed. The analysis process was carried out using RapidMiner software with stages including data preprocessing, normalization, and model evaluation using the Davies-Bouldin index. The results of the study indicate that the clustering approach is effective in identifying student behavior segmentation, which can then be used as a basis for formulating library service development policies in a more adaptive and targeted manner. This study contributes in the form of a data-based framework that can be replicated by other educational institutions to improve the effectiveness and efficiency of library management.

Keywords : Data Mining, K-Means, Clustering, Student Behavior, RapidMiner

1. PENDAHULUAN

Perpustakaan merupakan salah satu sarana strategis dalam mendukung keberhasilan proses pembelajaran dan riset di lingkungan perguruan tinggi. Keberadaan perpustakaan yang dikelola dengan baik mampu menjadi pusat sumber belajar yang efektif. Perpustakaan merupakan fasilitas yang mempunyai peran penting di perguruan tinggi. Perpustakaan selain berfungsi sebagai titik utama terkumpulnya informasi, perpustakaan juga digunakan sebagai tempat membaca dan mencari referensi serta tempat untuk meminjam buku bagi mahasiswa. [1]. Di era digital, pengambilan keputusan tanpa data sudah tidak lagi memadai. Pengambilan keputusan yang didasarkan pada analisis data cenderung menghasilkan strategi yang lebih tepat sasaran dibandingkan dengan pendekatan intuitif semata. Kemajuan teknologi informasi, pendekatan berbasis data semakin banyak digunakan termasuk pada sistem perpustakaan yang dimana diharapkan untuk memberikan layanan yang sesuai dengan kebutuhan pengguna. Salah satunya teknologi yang berperan penting yaitu Big data yaitu yang

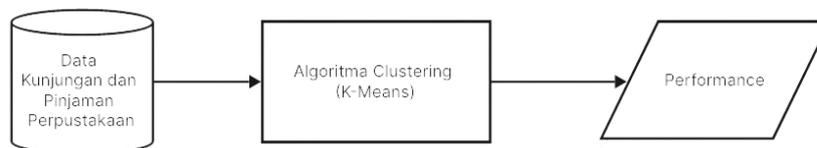
merujuk pada kumpulan data yang sangat besar dan kompleks, memungkinkan organisasi seperti perpustakaan untuk menganalisis informasi secara mendalam dan membuat keputusan yang lebih terinformasi. [2] Saat ini pengolahan big data semakin diminati karena kemudahannya dalam menggunakan. Dapat disimpulkan bahwa analisis data berbasis teknologi semakin banyak digunakan agar dapat membantu pengguna dalam mengambil keputusan strategis khususnya menyediakan koleksi buku di perpustakaan. [3] Dengan menggunakan big data dapat diolah menggunakan pendekatan data mining dengan metode clustering, untuk menggali pola-pola tersembunyi dari data perilaku pengguna perpustakaan.

Melalui pengelompokan berdasarkan frekuensi kunjungan dan jumlah peminjaman buku, pihak pengelola dapat memperoleh pemahaman yang lebih dalam tentang karakteristik mahasiswa serta menyusun strategi layanan yang lebih adaptif dan efisien. Sistem clustering pada data peminjaman buku dapat membantu perpustakaan mengelompokkan pengguna aktif, menengah, dan pasif, yang kemudian dapat dijadikan dasar dalam merancang strategi layanan.

Permasalahan utama yang diangkat dalam penelitian ini adalah belum adanya pemetaan perilaku mahasiswa dalam menggunakan layanan perpustakaan secara sistematis. Akibatnya, pendekatan dalam pengembangan fasilitas cenderung bersifat umum dan tidak mempertimbangkan keberagaman karakter pengguna. Belum tersedia pula sistem clustering berbasis data yang dapat membantu pihak perpustakaan mengenali kelompok pengguna aktif, sedang, maupun pasif. Penelitian ini analisis pola kunjungan dan jumlah peminjaman buku di perpustakaan universitas ma'soem menggunakan metode k-means. Penelitian ini dapat membantu pustakawan merancang layanan yang tepat sasaran.

2. METODE DAN BAHAN

Pada tahapan penelitian ini, prosedur kerja dilakukan seperti yang ditunjukkan dalam gambar berikut:



Gambar 1. Prosedur Kerja
Sumber: Rahayu & Dewi , 2025 [4]

2.1 Perpustakaan

Perpustakaan adalah ruangan, bagian sebuah gedung atau gedung itu sendiri yang digunakan untuk menyimpan buku dan terbitan lainnya, disimpan menurut tata susunan tertentu, digunakan untuk pembaca dan bukan untuk dijual. [5] Pengelolaan perpustakaan menjadi semakin kompleks. Dari sini awal mulai berkembang ilmu dan teknik mengelola perpustakaan. Perpustakaan sebagai sistem pengelolaan rekaman gagasan, pemikiran, pengalaman, dan pengetahuan umat manusia, mempunyai fungsi utama melestarikan hasil budaya umat manusia tersebut, khususnya yang berbentuk dokumen karya cetak dan karya rekam lainnya, serta menyampaikan gagasan, pemikiran, pengalaman, dan pengetahuan umat manusia itu kepada generasi-generasi selanjutnya. [6]

2.2 Data Mining

Data mining merupakan proses menemukan pola dan informasi penting dari kumpulan data yang besar. Tujuan dari data mining yaitu untuk mengubah data mentah menjadi wawasan yang dapat digunakan untuk membuat keputusan yang lebih baik. [7] K-Means clustering merupakan sebuah algoritma unsupervised learning yang digunakan dalam pengelompokan data dalam dataset yang tidak memiliki label kedalam sebuah cluster-cluster yang berbeda. [8]. Data mining biasa juga disebut dengan “Data atau knowledge discovery” atau menemukan pola tersembunyi pada data. [9] Knowledge Discovery In Database (KDD) adalah metode untuk memperoleh untuk memperoleh pengetahuan dari database yang terdapat tabel-tabel yang saling berelasi. Knowledge Discovery In Database (KDD) berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah kumpulan data. Hasil yang diperoleh dalam proses tersebut dapat digunakan sebagai basis pengetahuan (knowledge base) untuk.[10]

2.3 Algoritma K-Means

Cluster adalah gabungan dari beberapa data yang memiliki sesuatu yang sama dan memiliki beberapa perbedaan dengan sebagian data dari kelompok lain. Pengelompokan berlainan dengan klasifikasi, karena tidak memiliki variabel objek didalam pengelompokan. pengelompokan tidak berusaha guna mengklasifikasikan, memprediksi angka variabel objek. [11] K-means merupakan algoritma yang mengklasterisasikan berdasarkan

partisi dan melakukan clustering melalui proses iterasi berkelanjutan sampai dengan bertemu kondisi akhir, proses iterasi berhenti dan hasil clustering adalah output. [12]

2.4 Metode Clustering

Clustering data akan mengelompokkan objek yang paling dekat dimana terdapat kesamaan dengan objek lain, serta data yang akan di-cluster diambil secara acak atau random. [13] analisis Clustering dapat digunakan untuk memahami pola perilaku pengguna perpustakaan dan mengidentifikasi sumber daya yang paling sering digunakan, yang kemudian bisa menjadi dasar untuk pengambilan keputusan dalam mengoptimalkan koleksi. [14] Clustering merupakan salah satu teknik penting dalam data mining yang digunakan untuk mengelompokkan data menjadi subset yang serupa berdasarkan karakteristik atau pola tertentu. Tujuan utama dari clustering adalah untuk mengidentifikasi struktur yang tersembunyi dalam data, yang dapat membantu dalam pemahaman lebih lanjut tentang kelompok atau kategori yang ada di dalamnya. [15] Pernyataan ini mendukung konsep bahwa metode clustering bukan sekali jalan, melainkan melibatkan siklus perbaikan terus-menerus untuk mencapai pemetaan pola yang optimal.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

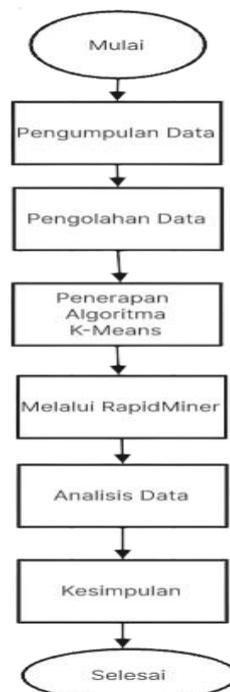
Untuk mendapatkan data-data yang dibutuhkan dalam penelitian ini, maka penulis menggunakan beberapa metode pengumpulan data diantaranya observasi dan pemintaan data kunjungan serta peminjaman buku kepada pustakawan Universitas Ma'soem.

3.2 Preprocessing Data

Pemrosesan data adalah langkah di mana data terpilih mengalami berbagai langkah pemrosesan untuk mempersiapkan data yang akan dianalisis. Data yang diperoleh kemudian diinput ulang ke dalam Microsoft Excel sebagai sarana pencatatan transaksi peminjaman buku, dan selanjutnya dilakukan proses pembersihan data (data cleaning) dengan menghapus entri yang tidak terbaca atau tidak valid guna memastikan akurasi dalam tahap analisis berikutnya. Fungsi utama pengolahan data adalah untuk memastikan bahwa data tersedia dan dapat diandalkan, serta meminimalkan kesalahan yang mungkin terjadi selama analisis, dalam proses ini data tersisa berjumlah 1.303 data.

3.3 Pemodelan

Algoritma K-Means adalah salah satu alat bantu yang dapat digunakan. Metode ini kemudian dapat mengolah dan menyederhanakan data atribut/numerik. Langkah-langkah yang dilakukan yang dapat dideskripsikan dalam gambar berikut :

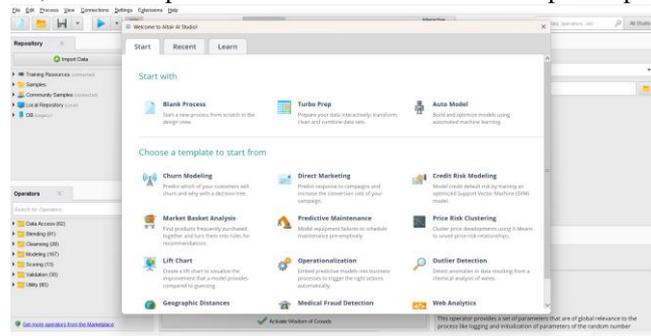


Gambar 2. Prosedur K-Means
Sumber: Rahayu & Dewi , 2025

3.4 mplementasi

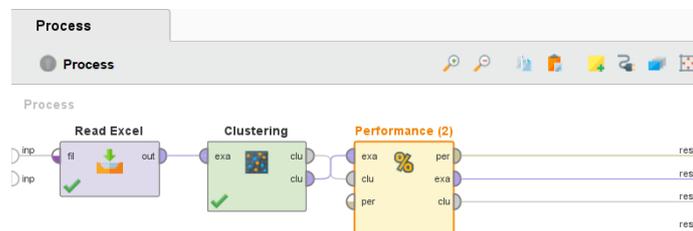
Menentukan Cluster dan Centroid

Buka software RapidMiner, kemudian pilih Blank Process untuk memulai proses pengujian.



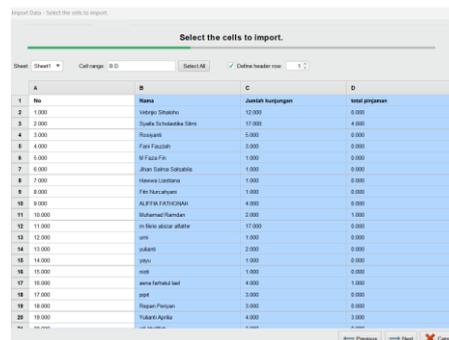
Gambar 3. Tampilan Awal RapidMiner

Kemudian tambahkan operator Read Excel untuk membawa file yang akan di import, K-Means untuk membentuk cluster dan operator Performance/Cluster Distance Performance untuk menampilkan hasil dari pembentukan cluster tersebut.



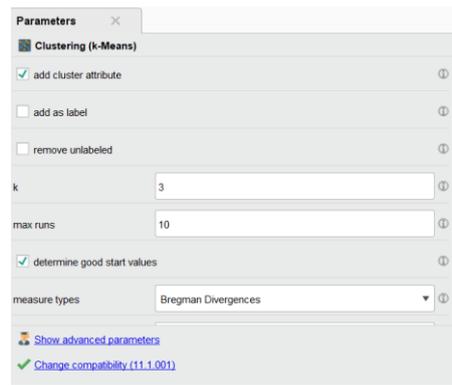
Gambar 4. Desain Operator K-Means

Import data yang akan digunakan pada Read Excel di klik 2 kali kemudian pilih file dan data yang akan digunakan dalam penelitian ini.



Gambar 5. Import Data

Pada operator Clustering atur jumlah cluster sesuai dengan cluster yang telah ditentukan sebelumnya yaitu 3. Jika sudah maka langsung klik tombol run.



Gambar 6. Penyesuaian Jumlah Cluster

Setelah proses sukses secara otomatis akan terlihat hasil cluster yang terbentuk dan informasi ringkas pada menu results.

Cluster Model

```
Cluster 0: 201 items
Cluster 1: 1072 items
Cluster 2: 30 items
Total number of items: 1303
```

Gambar 7. Hasil Cluster Model

Dari gambar di atas dapat dilihat bahwa cluster yang terbentuk oleh RapidMiner yaitu Cluster 0 berjumlah 201 data, Cluster 1 berjumlah 1072 data, dan Cluster 2 berjumlah 30 data. Kemudian pada Gambar 8 dapat dilihat nilai centroid yang secara otomatis dipilih untuk masing-masing clusternya.

Attribute	cluster_0	cluster_1	cluster_2
Jumlah kunjungan	5.269	1.509	13.900
total pinjaman	0.672	0.140	0.900

Gambar 8. Centroid Table

Pada halaman ExampleSet dapat terlihat data anggota masing- masing cluster yang terbentuk.

Row No.	id	cluster	Jumlah kunj...	total pinjaman
1	1	cluster_2	12	0
2	2	cluster_2	17	4
3	3	cluster_0	5	0
4	4	cluster_1	3	0
5	5	cluster_1	1	0
6	6	cluster_1	1	0
7	7	cluster_1	1	0
8	8	cluster_1	1	0
9	9	cluster_0	4	0
10	10	cluster_1	2	1
11	11	cluster_2	17	0
12	12	cluster_1	1	0
13	13	cluster_1	2	0
14	14	cluster_1	1	0
15	15	cluster_1	1	0
16	16	cluster_0	4	1
17	17	cluster_1	3	0
18	18	cluster_1	3	0

Gambar 9. Data Clustering

3.5 Validasi Performance

DBI merupakan metode untuk mengevaluasi kinerja pengelompokan suatu data. DBI digunakan sebagai matriks pengelompokan karena cara pengelompokan validasi untuk menilai kinerja hasil clustering berisi dua kategori yaitu validasi internal dan validasi eksternal. Adapun Langkah-langkah untuk menghitung nilai DBI adalah pertama peneliti harus menghitung Sum of Square Within-Cluster (SSW), menghitung sum of square between-cluster (SSB) kemudian baru dapat menghitung Davies Bouldin Index (DBI).

Menghitung Sum of Square Within-Cluster (SSW)

Data yang akan diuji dalam perhitungan SSW yaitu berjumlah 1.303 data dimana di ambil dari data pada iterasi terakhir akan disesuaikan dengan cluster data yang tersedia, sehingga menghasilkan sebagai berikut :

Tabel 1. Perhitungan SSW

0.7309998504	1.674630601	3.329405151
SSW0	SSW1	SSW2

Menghitung sum of square between-cluster (SSB)

Data yang akan diuji dalam perhitungan SSB yaitu berjumlah 1.303 data dimana di ambil dari data pada iterasi terakhir akan disesuaikan dengan cluster data yang tersedia, sehingga menghasilkan sebagai berikut :

Tabel 2. Perhitungan SSB

SSB(0,1)	3.886407603
SSB(0,2)	1.302170738
SSB(1,2)	9.146913456

Dari hasil di atas maka selanjutnya dapat menghitung validasi performance berupa Davies Bouldin Index (DBI) yang dideskripsikan pada halaman PerformanceVector. Perhitungan DBI menghasilkan sebagai berikut :

Tabel 3. Perhitungan DBI

DBI

0.2063285427

Nilai DBI dari proses clustering sebesar 0.206328, Nilai terkecil dari IDB menunjukkan jumlah cluster paling optimal. [16] itu berarti cluster yang terbentuk dinyatakan optimal karena cukup dekat dari nilai 0.

4. KESIMPULAN

Penelitian ini berhasil mengidentifikasi permasalahan utama dalam pengelolaan layanan perpustakaan, yaitu ketiadaan pemetaan perilaku mahasiswa secara sistematis yang menyebabkan pengembangan fasilitas bersifat umum dan kurang adaptif terhadap keberagaman pengguna. Melalui penerapan algoritma K-Means terhadap data kunjungan dan peminjaman buku selama 3 bulan terakhir dengan jumlah 1.303 data dan menghasilkan pengelompokan mahasiswa ke dalam klaster perilaku yang berbeda, yakni aktif, sedang, dan pasif. Data yang didapatkan bahwa kunjungan dan peminjaman buku di perpustakaan ma'soem dengan Cluster 0 berjumlah 201 data, Cluster 1 berjumlah 1072 data, dan Cluster 2 berjumlah 30 data. Kemudian nilai DBI dari proses clustering sebesar 0.206328, sehingga dinyatakan optimal karena cukup dekat dari nilai 0. Maka didapati bahwa kunjungan dan peminjaman buku di Universitas Ma'soem yakni aktif. Hasil klasterisasi ini tidak hanya memberikan pemahaman yang lebih dalam tentang karakteristik pengguna, tetapi juga menghasilkan rekomendasi berbasis data yang dapat dijadikan acuan strategis dalam pengambilan keputusan pengembangan layanan perpustakaan secara lebih efektif dan terarah. sekaligus memungkinkan pustakawan merancang layanan yang tepat sasaran berdasarkan kebutuhan mahasiswa.

REFERENSI

- [1] A. Azis, M. Ardiansah, H. Yanto, and M. Odi, "Peran Perpustakaan Dalam Memperkaya Ilmu Pengetahuan Mahasiswa Di Perguruan Tinggi," *J. Lentera Pendidik. Pus. Penelit. Lppm Um Metro*, vol. 9, no. 2, p. 270, 2024, doi: 10.24127/jlpp.v9i2.3859.
- [2] S. A. Ramadhana, C. A. Sukmana, R. Hidayat N, I. R. Kusumasari, and P. A. Bisnis, "Pengambilan Keputusan Berbasis Data Di Era Digital," *Triwikrama J. Ilmu Sos.*, vol. 5, no. 4, 2024.
- [3] L. Rosiana and I. Yuadi, "K-Means Clustering untuk Analisis Tren Peminjaman Buku di Perpustakaan," *J. Technol. Informatics*, vol. 7, no. 1, pp. 1–10, 2025, doi: 10.37802/joti.v7i1.933.
- [4] A. Rahayu and S. Dewi, "Klasterisasi Pemilihan Paket Umrah Berdasarkan Musim untuk Menentukan Strategi Promosi Menggunakan Algoritma K-Means," *J. Dimamu*, vol. 4, no. 1, pp. 83–96, 2025, doi: 10.32627/dimamu.v4i1.1326.

- [5] A. Nurohman, "Perpustakaan sebagai Teropong Profesionalisme Pustakawan," *Tik Ilmeu J. Ilmu Perpust. dan Inf.*, vol. 2, no. 1, p. 71, 2018, doi: 10.29240/tik.v2i1.389.
- [6] T. Kurniawan, "Peran Perpustakaan Perguruan Tinggi dalam Mengembangkan Institusional Repository," *Pustaloka*, vol. 8, no. 2, pp. 232–243, 2016, [Online]. Available: <https://jurnal.iainponorogo.ac.id/index.php/pustakaloka/article/view/683>
- [7] M. Nara Bagdja, R. Taufiq Subagio, and V. Dwi Kartika, "Penerapan Metode K-Means Untuk Mengklasifikasi Minat Konsumen Terhadap Produk Toko Online," *JATI (Jurnal Mhs. Tek. Inform., vol. 8, no. 5, pp. 10690–10695, 2024, doi: 10.36040/jati.v8i5.11144.*
- [8] A. Yudhistira and R. Andika, "Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering," vol. 1, 2023.
- [9] I. K. J. Arta, G. Indrawan, and G. Rasben Dantes, "Data Mining Rekomendasi Calon Mahasiswa Berprestasi di STMIK Denpasar Menggunakan Metode Technique For Other Reference By Similarity to Ideal Solution," *J. Ilmu Komput. Indones.*, vol. 4, no. 1, pp. 11–21, 2019.
- [10] H. Astuti, "Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Pengelompokan Data Pelanggan (Studi Kasus : PT. Pinus Merah Abadi)," *J. Web Inform. Teknol.*, vol. 4, no. 1, p. 9, 2020.
- [11] D. Nafila, V. Riyanto, T. Informatika, S. Informasi, U. Bina, and S. Informatika, "Penerapan Algoritma K-Means Clustering Pada Pola Kunjungan Perpustakaan menggunakan Soft System Methodology," vol. 13, 2024.
- [12] D. A. Fakhri, S. Defit, and Sumijan, "Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering," vol. 3, 2021.
- [13] A. S. P. Nisriina Nuur Hasanah, "Implementasi Data Mining Untuk Pengelompokan Buku Menggunakan Algoritma K-Means Clustering," vol. 4, no. 2, pp. 300–311, 2022.
- [14] A. F. Zabidi, "Penerapan Algoritma K-Means untuk Pengelompokan Koleksi Perpustakaan dengan Data Mining," *Media J. Inform.*, vol. 16, no. 2, p. 233, 2024, doi: 10.35194/mji.v16i2.4814.
- [15] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *J. Ilm. Inform. dan Ilmu Komput.*, vol. 3, no. 1, pp. 46–56, 2024, doi: 10.58602/jima-ilkom.v3i1.26.
- [16] E. Muningsih, I. Maryani, and V. R. Handayani, "Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa," *J. Sains dan Manaj.*, vol. 9, no. 1, pp. 95–100, 2021, doi: 10.31294/evolusi.v9i1.10428.