

## MODEL KLASTERISASI DAN ANALISIS SISTEM REKOMENDASI KEPADA PENGGUNA PERPUSTAKAAN UNIVERSITAS 17 AGUSTUS 1945 SURABAYA MENGGUNAKAN K-MEANS

Mochammad Rizki Firmansyah<sup>1</sup>, Ahmad Habib<sup>2</sup>

<sup>1,2</sup>Universitas 17 Agustus 1945 Surabaya, Indonesia

[Firmansyahrizki000@gmail.com](mailto:Firmansyahrizki000@gmail.com)<sup>1\*</sup>

[habib@untag-sby.ac.id](mailto:habib@untag-sby.ac.id)<sup>2</sup>

Received: 22-09- 2024

Revised: 22-10-2024

Approved: 20-11-2024

### ABSTRAK

*Dalam situasi ini, analisis data menjadi alat yang sangat penting untuk memahami minat membaca pengguna dan mengukur minat mereka terhadap sistem rekomendasi perpustakaan. Untuk menganalisis minat membaca pengguna, algoritma pembelajaran mesin K-means dapat digunakan untuk memberikan informasi yang lebih mendalam dan efektif. Tujuan penelitian ini adalah mengembangkan model clustering menggunakan K-Means untuk menganalisis dan mengklusterkan sistem rekomendasi bagi pengguna perpustakaan melalui top 3 buku pada setiap prodi dengan mengelompokkannya berdasarkan 3 digit pertama NIM (Nomor Induk Mahasiswa). K-Means adalah algoritma clustering yang umum dan mudah dipahami. K-Means clustering adalah algoritma partisi data mining yang populer dan mudah digunakan untuk mengelompokkan data yang tidak berlabel ke dalam kelompok (cluster) berdasarkan kesamaan. Algoritma K-Means bekerja dengan mengelompokkan data berdasarkan jarak ke centroid dan menghitung ulang centroid berdasarkan cluster yang telah terbentuk. Hasil yang didapat dari pengujian yang telah dilakukan dengan metode K-Means adalah didapatkan 16 jumlah cluster berdasarkan pengelompokkan NIM (Nomor Induk Mahasiswa), dari pengelompokkan NIM (Nomor Induk Mahasiswa) tersebut menghasilkan rekomendasi buku berupa 3 judul buku dengan frekuensi peminjaman pada setiap pengelompokannya atau clusternya.*

**Kata kunci:** K-Means, Clustering, Perpustakaan

### PENDAHULUAN

Perpustakaan merupakan salah satu hal yang menunjang pendidikan, penelitian dan pengetahuan di masyarakat. Pemanfaatan perpustakaan oleh berbagai kalangan, seperti mahasiswa, peneliti, dan masyarakat umum, mempunyai dampak yang signifikan terhadap efisiensi dan keberlanjutan perpustakaan itu sendiri. Oleh karena itu, penting untuk memahami dan mengukur minat baca para pengguna perpustakaan.

Mengukur permintaan pengguna penting untuk mengidentifikasi kelemahan dan kelebihan layanan perpustakaan. Untuk memenuhi kebutuhan pengguna, perpustakaan harus terus mengembangkan dan meningkatkan layanannya untuk memenuhi harapan pengguna. Dalam konteks ini, analisis data menjadi alat yang sangat penting untuk memahami apa yang menjadi minat baca para pengguna dan mengukur tingkat minat baca mereka pada sistem rekomendasi perpustakaan. Penggunaan algoritma pembelajaran mesin, seperti K-Means, dalam analisis minat baca pengguna dapat memberikan wawasan yang lebih dalam dan efektif. Metode K-Means adalah salah satu algoritma pengelompokan yang populer dalam machine learning. [1].

Oleh karena itu, Clustering merupakan teknik pengelompokan data berdasarkan kemiripan dalam satu cluster. [2] [3] Penelitian ini bertujuan untuk mengembangkan model klasterisasi menggunakan K-Means untuk menganalisis dan klusterisasi sistem rekomendasi pengguna perpustakaan. Penelitian ini dapat dijadikan sebagai acuan

rekomendasi untuk pengoptimalisasikan pelayanan perpustakaan baik untuk tata letak maupun pengadaan buku dengan memprioritaskan jenis buku yang sangat diminati. [4] [5] [6] Dengan model ini diharapkan akan lebih mudah untuk mengidentifikasi buku yang paling diminati terhadap perpustakaan untag Surabaya. [7] Hasil analisis ini dapat memberikan landasan untuk meningkatkan layanan perpustakaan dan meningkatkan tingkat baca kepada para pengguna perpustakaan untag Surabaya melalui rekomendasi buku. [8] K-Means merupakan suatu metode non hierarki, pada tahap proses K-Means pusat cluster di pilih secara acak dari sekumpulan data komponen di dalam populasi data dan menandai komponen tersebut ke salah satu pusat cluster yang telah didefinisikan tergantung dari jarak minimum antar komponen dengan tiap-tiap cluster [9]

Penelitian ini akan memberikan kontribusi penting dalam bidang informasi dan ilmu perpustakaan serta dapat digunakan oleh berbagai jenis perpustakaan, mulai dari perpustakaan akademik hingga perpustakaan umum. Hasil yang didapatkan untuk data kategori buku yang berada pada cluster paling diminati, inilah nantinya yang akan dijadikan sebagai bahan evaluasi bagi pihak pustakawan dalam meningkatkan koleksi buku diperpustakaan [10] Dengan memahamiminat baca pengguna, perpustakaan dapat bekerja lebih efektif untuk memenuhi kebutuhan pengguna dan meningkatkan pengalaman mereka.

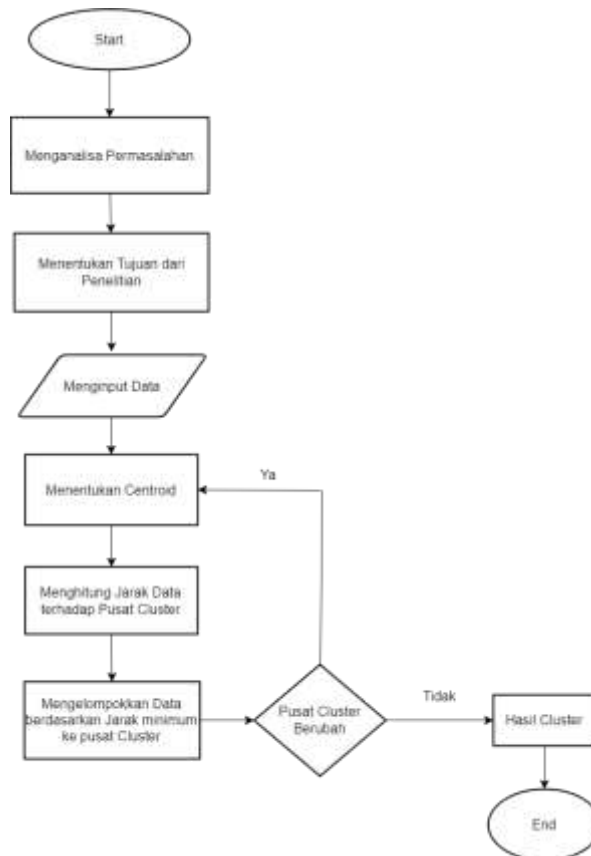
## **METODE PENELITIAN**

Metode penelitian ini menggunakan K-Means clustering yang dimana metode ini menjelaskan secara sistematis dan berguna sebagai pedoman untuk melakukan penelitian. Mulai dari pengumpulan data dan bahan penelitian yang sesuai dengan survei yang telah dilakukan. Data yang diambil adalah data dari peminjam buku di perpustakaan. Tujuan dari penelitian ini diharapkan agar bisa berjalan dengan baik sesuai dengan tujuan awal dari penelitian ini. Agar mempermudah dalam pemahaman menyelesaikan permasalahan yang dihadapi, sehingga melakukan penyusunan kerangka secara teratur dan jelas. Dan untuk tools yang digunakan adalah Excel dan Python.

Kegunaan dari excel adalah untuk mengubah format data dari excel ke csv dan menjadikan data selama beberapa bulan menjadi satu agar dapat diolah pada python ataupun google colab.

Dan untuk Python atau google colab digunakan untuk mengklusterisasi data dengan memasukkan metode k-means agar dapat menemukan jumlah clusternya pada setiap data Nomor Induk Mahasiswa, lalu langkah selanjutnya adalah menentukan top 3 buku dengan jumlah peminjaman pada setiap bukunya yang nantinya akan menampilkan hasil akhir adalah setiap Nomor Induk Mahasiswa memiliki perkomendasi buku tersendiri.

Metode yang dapat digunakan adalah metode K-Means dan yang dapat digunakan untuk mengelompokkan data dari beberapa sampel yang didapat dari dataset para peminjam buku di perpustakaan. Hasil klusterisasi akan dievaluasi performanya untuk bahan pertimbangan langkah selanjutnya. Berikut tahapannya:



Gambar 1 Tahapan Penelitian

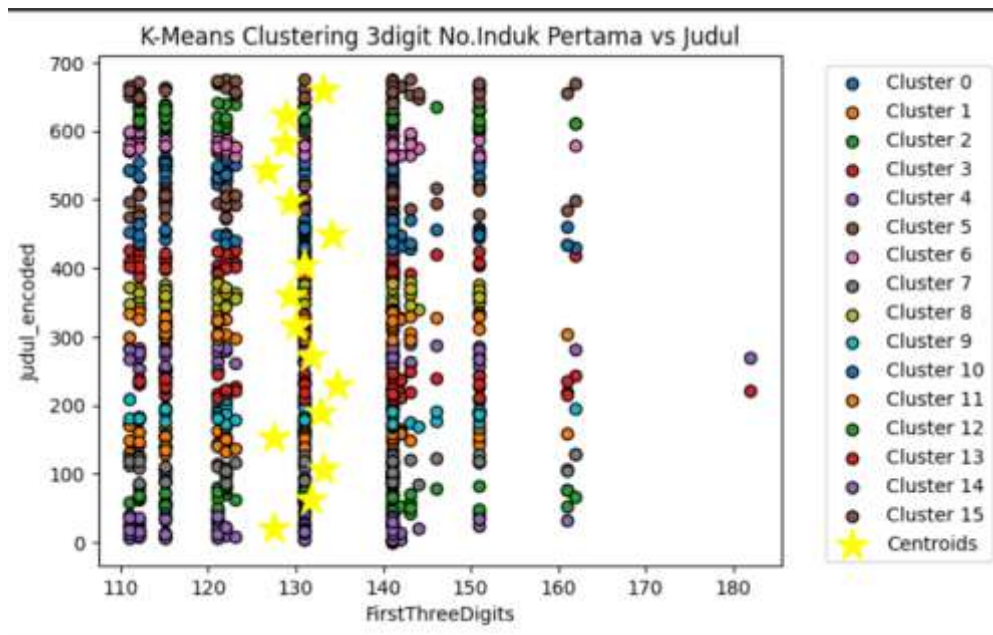
## HASIL DAN PEMBAHASAN

Pada tahapan pengujian yang digunakan untuk proses clustering adalah menggunakan K-Means yang dimana hal tersebut akan mempermudah proses clustering pada data. Mulai dari tahap menginput dataset perpustakaan sampai mendapatkan hasil rekomendasi 3 buku pada setiap NIM (Nomor Induk Mahasiswa).

```
Nilai centroid per kluster:  
Cluster 1: [126.51428571 540.95714286]  
Cluster 2: [131.91860465 108.98837209]  
Cluster 3: [130.32467532 319.44155844]  
Cluster 4: [136.90909091 202.59090909]  
Cluster 5: [129.07042254 621.22535211]  
Cluster 6: [134.06756757 449.2027027 ]  
Cluster 7: [127.41333333 19.53333333]  
Cluster 8: [129.43333333 239.35 ]  
Cluster 9: [128.11111111 159.45555556]  
Cluster 10: [130.74666667 405.45333333]  
Cluster 11: [129.30985915 581.18309859]  
Cluster 12: [134.91803279 276.21311475]  
Cluster 13: [129.08108108 495.83783784]  
Cluster 14: [132.11111111 62.37037037]  
Cluster 15: [132.56896552 658.81034483]  
Cluster 16: [128.61111111 361.47222222]
```

Gambar 2 Hasil centroid per cluster

Setelah melakukan clustering terdapat 16 cluster, yang dimana berfungsi sebagai pusat massa setiap cluster. Mereka memberikan ringkasan numerik karakteristik unik titik data dalam cluster tertentu. Jumlah cluster yang telah ditentukan adalah sebanyak 16 cluster.



Gambar 3 Hasil Clustering

Menampilkan visualisasi hasil clustering. Didapatkan sebanyak 16 cluster berdasarkan NIM yang dimana setiap cluster memiliki warna yang berbeda agar memudahkan dalam menganalisa setiap cluster dan bentuk bintang berwarna kuning menunjukkan centroidnya.

```
[63] import pandas as pd

keterangan_mapping = [
    '111': 'Adm Negara',
    '112': 'Adm Niaga',
    '113': 'Tikone',
    '121': 'Manajemen',
    '122': 'Akuntansi',
    '123': 'Ekonomi Pembangunan',
    '131': 'F. Hukum',
    '141': 'FT Industri',
    '142': 'FT Mesin',
    '143': 'FT Sipil',
    '144': 'FT Arsitek',
    '146': 'FT Informatika',
    '151': 'Psikolog',
    '161': 'B. Inggris',
    '163': 'B. Jepang',
    '182': 'Teknologi Listrik',
]

# Menyiapkan Dataframe data untuk mapping keterangan
data1 = pd.DataFrame({
    'FirstThreeDigits': list(keterangan_mapping.keys()),
    'keterangan': list(keterangan_mapping.values())
})

# Menampilkan Dataframe data yang telah dibuat
data1
```

Gambar 4 Keterangan dari 16 cluster

Menyiapkan dataframe data untuk mapping keterangan agar menampilkan setiap Prodi sesuai dengan Nomor Induk Mahasiswa yang telah dilakukan clustering. Terakhir menampilkan dataframe yang telah dibuat, yang menyajikan data mapping antara 3 digit pertama pada NIM dan nama jurusan dalam format tabel.

	Judul	Cluster	FrekuensiPeminjaman	Keterangan
0	Tutugel - Cikal Bakal Majalah / Hermas Sup...	0	2	Ekonomi Pembangunan
1	Personal Branding Format Baru - Membangun Cb...	0	2	Ekonomi Pembangunan
2	123 Tips Sempai Cerdik Awak Muda dan Sehat dan...	0	3	Ekonomi Pembangunan
3	Jalan Indah Para Raksas Allah - Menggapai Pa...	1	2	FT Sipi
4	Selebaran Press / Alexander Puchta	1	2	FT Sipi
5	Andi Seokigunya, Cikan / Wahyu Aditya & Titik	1	1	FT Sipi
6	Namku Islamku Islam Kita - Dewa dan Islam B...	2	2	Adm Naga
7	Soul Medals - Kisah-Kisah Keberanian dan Kebal...	2	2	Adm Naga
8	Lancar Membaca Horat dan Angka / Kak Adi	2	2	Adm Naga
9	Seri Bericara Pada Anak - Panduan Mendidik An...	3	2	Psikolog
10	Fikah Komunikasi - Tradisi, Teori, dan Meta...	3	2	Psikolog
11	Ibu, Gedung Apa? / EDAH Iyobony	3	2	Psikolog
12	Deskat Bato - Kabupaten Karanganyar / Di...	4	3	F Hukum
13	Pembentayan Usaha Kecil Menengah - Upaya Memb...	4	2	F Hukum
14	Sepanjang Mata Kelabu dan Harung-Burung yang Ma...	4	2	F Hukum
15	Kata Bahasa dan Sastra - Kata Metafora Le...	5	2	Ikoni
16	Hamletika - Seri Khutbah & Hikmah	5	2	Ikoni
17	Pasar Modal Syariah - Mengenal dan Memahami Re...	5	2	Ikoni
18	Bahasa Klat Bahasa Korea - Tata Bahasa, Ejaan...	5	2	B Inggris
19	New Normal Covid - Panduan Menjalani Tahapan H...	6	2	B Inggris
20	Agar Sekolah Ditemu Berbuah Kelayakan - Inilah...	6	1	B Inggris
21	100 Hal yang Tidak Kalian Ketahui Tentang Saia...	7	2	FT Industri
22	Pendidikan Pancasila / Muhammad Maza Adha	7	2	FT Industri

Gambar 5 Menampilkan top 3 buku pada setiap prodi

23	Hanya, Tak Ada yang Dapat Bersekolah di Arta...	7	3	FT Industri
24	Honor (Sampul) Pangkal Kaya / Dora Isdanti	8	1	Teknologi Liris
25	Jari Tengah / Adlan Djaparaling	8	1	Teknologi Liris
26	Kisah-Kisah Kecil & Gambil / Agus Moor	9	2	Manajemen
27	Sistem Simbol Budaya Indonesia	9	2	Manajemen
28	Rambutan Tutugelan di Wajolito / Tora Uyo	9	2	Manajemen
29	Berkarya dan Perilaku Sosial Gaya Generasi Milen...	10	1	FT Informatika
30	Caracari Si Gila - Kumpulan Puisi / Muhammad Al...	10	1	FT Informatika
31	Enklipedi Orang Kudu Sepanjang Tahun	10	1	FT Informatika
32	Roong Dessert Pilihan untuk Hiliris Doge	11	2	Akuntansi
33	Buku Catatan Maya / Isabel Alenda	11	2	Akuntansi
34	Rudhas Koneksi / Ardy Krista Oenata	11	2	Akuntansi
35	Citra Sepanjang Masa / Kiki Gibran	12	2	Adm Negera
36	Surat Pribat Elab / Riya Dwikhangkudhoro	12	2	Adm Negera
37	Berg Pembenerak / Kiki Gibran	12	2	Adm Negera
38	Home Slavy For Fast Study - Anyk Belajar Bah...	13	2	FT Meas
39	181 Dos Mutan	13	1	FT Meas
40	36 Praktik Baik Literasi Sekolah - Mewujudkan ...	13	1	FT Meas
41	Cikan & Teman Rie Sogila / Wahyu Aditya & In...	14	2	B Jepang
42	Selakuon - Nivona Pemyantahan 9 Hari Berba B...	14	2	B Jepang
43	Baba Barani Berjendapat / Citra Auliana	14	1	B Jepang
44	40+ Kisah Makkazi Dalam Alkitab / Andras Rad...	15	1	FT Arabik
45	Big Book Belajar Matematika / Coomas Yuzkarto	15	1	FT Arabik
46	Dangang Nisa Baba - Pustaka Dangang Pempantar	15	1	FT Arabik

Gambar 6 Menampilkan top 3 buku pada setiap prodi

Menampilkan hasil dari dataframe baru yang dimana semua buku-buku yang paling banyak dipinjam dan dijadikan sebagai bahan rekomendasi dan juga menampilkan dari keseluruhan top 3 buku berdasarkan judul, cluster ke berapa, frekuensi peminjaman, dan juga keterangan jurusan atau prodi.

## KESIMPULAN

Berdasarkan penelitian yang telah dilaksanakan, dimulai dari mencari dataset hingga menentukan hasil clustering, perancangan tahapan urut dari awal pencarian dataset hingga akhirnya mendapatkan beberapa cluster dan menentukan centroidnya. Penerapan K-Means Clustering digunakan sebagai metode untuk merekomendasikan buku kepada pengguna perpustakaan sesuai dengan keilmuan atau kategori. Kesimpulan dari

sistem rekomendasi ini berupa:

Penentuan Jumlah Cluster dalam K-Means:

1. Cluster yang didapatkan sebanyak 16 cluster, dikelompokkan berdasarkan NIM (Nomor Induk Mahasiswa).
2. Setiap cluster memiliki tingkat rekomendasi buku sebanyak 3 buku yang dibedakan berdasarkan tingkat peminjaman buku pada setiap cluster.
3. Didapatkan hasil rata-rata dari frekuensi peminjaman adalah sebanyak 3x, 2x, 1x.

Contoh dari penjelasan sistem:

Apabila ada mahasiswa teknik informatika yang ingin meminjam buku dengan kategori komputer maka akan menampilkan 3 buku teratas sesuai dengan kategori yang diinginkan begitupun dengan jurusan yang lain. Hasil yang didapat adalah perekomendasi hanya kepada mahasiswa belum termasuk dosen dan pengurus universitas.

## **DAFTAR PUSTAKA**

Penulisan daftar pustaka menggunakan Format kutipan IEEE. Harus menggunakan tools seperti Mendeley, Zotero maupun reference management tools yang lain.

## **DAFTAR PUSTAKA**

- [1] N. R. Fitradhi, M. Firman Hidayat, T. W. Saputro, M. G. Alifian, and A. P. Sari, "LAPORAN FINAL PROJECT MACHINE LEARNING REKOMENDASI MUSIK SPOTIFY MENGGUNAKAN METODE K-MEANS," 2023.
- [2] H. Tommy Argo Simanjuntak, P. Ephraim Prabowo Silaban, J. Koko Sarasi Manurung, and V. Handayani Sormin, "KLAUSTERISASI BERITA BAHASA INDONESIA DENGAN MENGGUNAKAN K-MEANS DAN WORD EMBEDDING," vol. 10, no. 3, pp. 641–652, 2023, doi: 10.25126/jtiik.2023106468.
- [3] R. Anjariansyah and A. Triayudi, "Clustering Kebutuhan Makanan untuk Meminimasi Standar Deviasi Angka Kebutuhan Gizi Menggunakan Algoritma K-Means dan K-Medoids," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 1, p. 597, Jan. 2022, doi: 10.30865/mib.v6i1.3522.
- [4] D. A. Fakhri, S. Defit, and Sumijan, "Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering," *Jurnal Informasi dan Teknologi*, pp. 160–166, Sep. 2021, doi: 10.37034/jidt.v3i3.137.
- [5] H. Mutiasari, T. W. Purboyo, and R. A. Nugrahaeni, "SISTEM REKOMENDASI FILM MENGGUNAKAN METODE K-MEANS CLUSTERING (MOVIE RECOMMENDATION SYSTEM USING K-MEANS CLUSTERING METHOD)."
- [6] U. Laili Musyarofah, S. Nur Alima, and D. K. Satria Yudha, "Prosiding Seminar Nasional Teknologi dan Sistem Informasi (SITASI) 2022 CLUSTERING METHOD." [Online]. Available: <http://sitasi.upnjatim.ac.id/>
- [7] 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.
- [8] D. Siburian, S. Retno Andani, I. Purnama Sari, and G. Artikel, "Implementasi Algoritma K-Means untuk Pengelompokan Peminjaman Buku Pada Perpustakaan Sekolah Implementation of K-Means Algorithm for Clustering Books Borrowing in School Libraries," *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, vol. 1, no. 2, pp. 2828–9099, 2022, doi: 10.55123/jomlai.v1i2.725.
- [9] A. Febrianto, S. Achmadi, and A. P. Sasmito, "PENERAPAN METODE K-MEANS UNTUK CLUSTERING PENGUNJUNG PERPUSTAKAAN ITN MALANG," 2021.
- [10] "Implementasi Data Mining Untuk Pengelempokan Buku Di Perpustakaan Yayasan Nurul

Islam Indonesia Baru Dengan Metode K-Means Clustering Haryani \* , Dicky Nofriansyah \*\* , Ita Mariami \*\* \* Program Studi Sistem Informasi, STMIK Triguna Dharma \*\* Program Studi Sistem Informasi, STMIK Triguna Dharma,” 2021, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jct/index>