

Contents list available at www.jurnal.unimed.ac.id

CESS
(Journal of Computing Engineering, System and Science)

journal homepage: <https://jurnal.unimed.ac.id/2012/index.php/cess>



**Rekomendasi Penambahan Koleksi Buku Menggunakan Algoritma K-Means
Clustering di Perpustakaan UINSU**

***Recommendations for Additional Book Collections Using the K-Means
Clustering Algorithm in the UINSU Library***

Dhea Aulia Nurhasanah^{1*}, Raissa Amanda Putri²

*^{1,2} Fakultas Sains dan Teknologi, Prodi Ilmu Komputer, Universitas Islam Negeri Sumatera Utara
Jl. Lap. Golf, Kp. Tengah, Kec. Pancur Batu, Kabupaten Deli Serdang, Sumatera Utara, Indonesia
email: ¹dheaauliaa21@gmail.com, ²raissa.ap@uinsu.ac.id*

ABSTRAK

Perpustakaan Universitas Islam Negeri Sumatera Utara (UINSU) merupakan salah satu perpustakaan yang berfungsi sebagai sarana untuk menyimpan bahan pustaka yang dipakai oleh mahasiswa untuk menggali ilmu sumber informasi serta menopang kegiatan studi di lingkungan perguruan tinggi negeri tersebut. Pada perpustakaan, sejumlah data dapat diperoleh berdasarkan data historis, sehingga data akan bertambah secara terus menerus, misalnya data transaksi peminjaman buku. Kesulitan dalam menentukan jenis koleksi buku apa yang seharusnya perlu menjadi prioritas untuk diperbanyak atau tidak harus menjadi pengetahuan penting bagi para pengelola perpustakaan tersebut karena, jika jumlah buku yang dipinjam tidak sebanding dengan stok buku yang ada, akan berpengaruh pada kurangnya minat baca mahasiswa untuk membaca. Untuk mengatasi permasalahan tersebut, maka perlu dibuat suatu teknik data mining menggunakan metode k-means *clustering* yang dapat membantu dalam pengelompokan buku yang sering dipinjam agar petugas perpustakaan mengetahui prioritas penambahan koleksi buku di perpustakaan tersebut. Pengelompokan ini didasarkan pada atribut yang digunakan, yaitu judul buku, jumlah buku yang dipinjam, dan waktu peminjaman. Hasil dari metode tersebut akan mengelompokkan data menjadi 2 kelas dimana *Cluster 0* untuk rekomendasi menambah jumlah stok buku, dan *cluster 1* tidak rekomendasi menambah jumlah stok buku.

Kata Kunci: *data mining; metode K-Means Clustering; Perpustakaan UIN Sumatera Utara (UINSU); koleksi buku.*

ABSTRACT

The North Sumatra State Islamic University (UINSU) library is a library that functions as a means of storing library materials used by students to explore information sources and

*Penulis Korespondensi:
email: dheaauliaa21@gmail.com

support study activities within the state university environment. In libraries, a number of data can be obtained based on historical data, so that data will increase continuously, for example book loan transaction data. The difficulty in determining what type of book collection should be a priority for expansion or not should be important knowledge for library managers because, if the number of books borrowed is not proportional to the existing stock of books, this will result in students' lack of interest in reading. To overcome this problem, it is necessary to create a data mining technique using the k-means *clustering* method which can help in grouping books that are frequently borrowed so that library staff know the priorities for adding book collections in the library. This grouping is based on the attributes used, namely the title of the book, the number of books borrowed, and the time borrowed. The results of this method will group the data into 2 classes where *Cluster 0* is for recommendations to increase the number of book stocks, and *cluster 1* is not for recommendations to increase the number of book stocks.

Keywords: *data mining; K-Means Clustering method; UIN Sumatera Utara Library (UINSU); book collection.*

1. PENDAHULUAN

Kemajuan teknologi saat ini berpengaruh pesat termasuk dalam pengolahan data khususnya dalam bidang perpustakaan yang mengharuskan menyimpan data arsip yang jumlahnya sangat banyak. Data *mining* adalah ilmu menambang data dalam database yang besar untuk mendapatkan informasi yang berharga dan berguna yang saat ini banyak dipakai dalam pengolahan data khususnya dalam bidang perpustakaan[1]. Perpustakaan Universitas Islam Negeri Sumatera Utara (UINSU) merupakan salah satu perpustakaan perguruan tinggi negeri yang berfungsi sebagai sarana sumber informasi dan ilmu pengetahuan untuk menyimpan bahan pustaka yang dipakai oleh mahasiswa untuk menggali ilmu sumber informasi serta menopang kegiatan studi di lingkungan perguruan tinggi negeri tersebut. Pada perpustakaan perguruan tinggi, sejumlah data dapat diperoleh berdasarkan data historis, sehingga data akan bertambah secara terus menerus, misalnya data transaksi peminjaman buku. Seperti halnya pada proses peminjaman buku, dalam sebuah perpustakaan perguruan tinggi akan menghasilkan data-data baru berupa data transaksi peminjaman buku oleh mahasiswa. Apabila dilakukan pengolahan data pada sumber data tersebut maka dapat diketahui berbagai informasi yang bermanfaat dalam membantu menentukan pengadaan koleksi perpustakaan[2].

Kesulitan dalam menentukan jenis koleksi apa yang seharusnya perlu menjadi prioritas untuk diperbanyak atau tidak, harus menjadi pengetahuan penting bagi para pengelola perpustakaan tersebut. Karena jika jumlah buku yang dipinjam tidak sebanding dengan stok buku yang ada, akan berpengaruh pada kurangnya minat baca mahasiswa untuk membaca. Pengelola perpustakaan harus memiliki pengetahuan berdasarkan data pengunjung, sehingga tidak terjadi kesalahan dalam pengadaan.

Untuk mengatasi permasalahan tersebut, maka perlu dibuat suatu teknik data *mining* menggunakan metode k-means *clustering* yang dapat membantu dalam pengelompokan buku yang sering dipinjam agar petugas perpustakaan mengetahui prioritas penambahan koleksi buku di perpustakaan tersebut. Dengan metode K- Means dapat berpartisipasi data ke dalam bentuk dua kelompok ataupun lebih. Metode tersebut akan membagi data ke dalam suatu kelompok dimana data yang ber karakteristik sama akan diletakkan ke dalam satu kelompok

yang sama sedangkan data yang memiliki karakteristik yang berbeda akan dikelompokkan ke dalam kelompok lainnya[3].

Penelitian sebelumnya dilakukan oleh Hasni Pratiwi (2022) dengan judul “Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means *Clustering*” yang membahas tentang penerapan suatu metode data mining dengan cara menganalisa data penjualan untuk menghasilkan informasi yang dapat dijadikan sebagai perencanaan dan pengendalian persediaan barang. Adapun pengolahan datanya dapat dilakukan melalui proses *clustering* data dengan menerapkan metode k-means *clustering*. Metode k-means *clustering* bertujuan mengelompokkan data yang mempunyai karakteristik yang sama ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam *cluster* lain. Dengan dilakukan pengelompokan data tersebut dapat diketahui barang apa saja yang sudah terjual dan barang apa saja yang masih ada dalam stok[4]. Berdasarkan hal tersebut peneliti akan membuat sebuah penelitian dengan metode K-Means *clustering* tersebut untuk rekomendasi penambahan koleksi buku di perpustakaan UINSU sebab dinilai lebih efektif dan efisien sehingga dapat membantu dalam tersedianya penambahan koleksi buku di perpustakaan tersebut.

2. LANDASAN TEORI

2.1. Data Mining

Data *mining* didefinisikan sebagai sebuah proses untuk menemukan hubungan, pola dan tren baru yang bermakna dengan menyaring data yang sangat besar yang tersimpan dalam penyimpanan, menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. Data *mining*, sering juga disebut sebagai *knowledge discovery in database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar[5].

2.2. Algoritma K-Means *Clustering*

K-Means merupakan algoritma *clustering* yang berulang-ulang. Algoritma k-means menetapkan nilai-nilai *cluster* (k) secara acak, dimana nilai tersebut menjadi pusat dari *cluster* atau disebut sebagai centroid, mean atau means. Pada algoritma pembelajaran ini, komputer melakukan kelompok sendiri data-data yang menjadi masukannya tanpa mengetahui terlebih dulu target kelasnya. Pembelajaran ini termasuk unsupervised learning. Dengan partitioning secara iteratif, K-Means mampu meminimalkan rata-rata jarak setiap data ke klasternya[6]. Tujuan utama *clustering* biasanya adalah mencari prototipe kelompok yang paling representatif terhadap data, memberikan abstraksi dari setiap obyek data dalam kelompok dimana sebuah data terletak di dalamnya. bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal [7]. Berikut merupakan rumus algoritma K-Means :

$$d_{x,y} = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$$

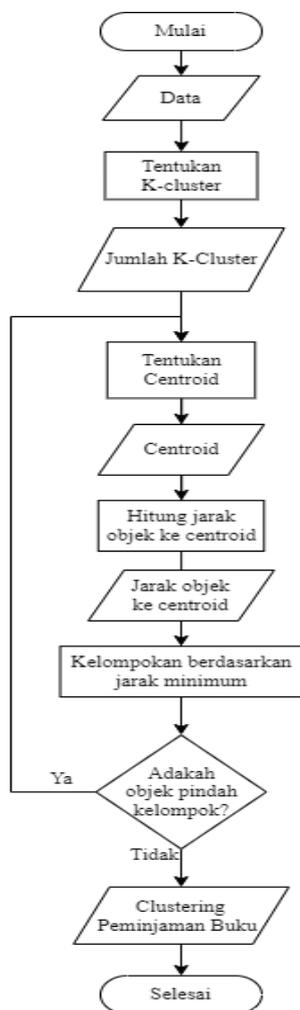
Algoritma ini umumnya menerima masukan berupa data tanpa label kelas. Hal ini berbeda dengan supervised learning yang menerima masukan berupa vektor (x1, y1), (x2, y2), ..., (xi, yi), di mana xi merupakan data dari suatu data pelatihan dan yi merupakan label kelas untuk xi.

2.3. Rapid Miner

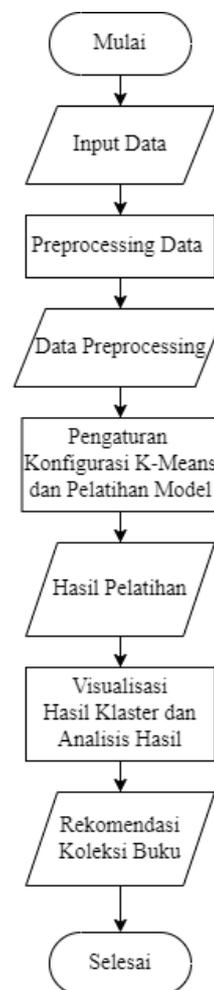
RapidMiner sebuah perangkat lunak ilmu yang dikembangkan oleh perusahaan yang bersama-sama menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, dan analisis *predictive*. RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data *mining*, *text mining* dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner digunakan dalam bidang bisnis, selain itu juga dalam bidang pendidikan, penelitian, pelatihan, *prototype* yang mendukung proses pembelajaran mesin seperti visualisasi, model, persiapan data dan optimasi [8].

3. METODE

Metode perancangan yang dipilih untuk penelitian ini adalah menggunakan flowchart. Dalam tahap ini dari proses Rekomendasi Penambahan Koleksi Buku Menggunakan Algoritma K-Means *Clustering* di Perpustakaan UINSU, langkah pentingnya adalah merancang aplikasi yang akan diimplementasikan. Aplikasi yang sedang dikembangkan memiliki langkah-langkah yang harus dijalani. Selanjutnya, langkah perancangan ini merupakan upaya untuk memahami bagaimana sistem akan dibangun untuk memenuhi kebutuhan dan bagaimana sistem tersebut akan beroperasi.



Gambar 1. Flowchart K-Means



Gambar 2. Flowchart Sistem RapidMiner

Flowchart K-Mean dimulai penentuan jumlah *cluster* (K) dan pemilihan centroid awal. Selanjutnya, perhitungan jarak antara setiap objek data dan centroid dilakukan, dan objek data ditempatkan dalam *cluster* dengan jarak minimum. Jika ada objek yang berpindah kelompok, iterasi dilakukan dengan menentukan centroid baru dan menghitung kembali jarak objek data ke centroid. Proses ini berlanjut sampai tidak ada perubahan *cluster*, yang menandakan konvergensi dan flowchart selesai.

Flowchart sistem RapidMiner dimulai dengan menyediakan dataset peminjaman buku dari perpustakaan UINSU sebagai input. Proses pre-processing data kemudian dilakukan untuk membersihkan dan mempersiapkan data. Selanjutnya, konfigurasi algoritma K-Means dilakukan dengan menentukan parameter seperti jumlah *cluster* (K). Setelah itu, model dilatih menggunakan operator K-Means pada RapidMiner. Hasil *cluster* dan posisi centroid kemudian divisualisasikan untuk memudahkan analisis. Proses analisis hasil dilakukan untuk memahami pola peminjaman buku di perpustakaan UINSU. Berdasarkan analisis *cluster*, rekomendasi penambahan koleksi buku disusun. Hasil rekomendasi disimpan atau disajikan sebagai output dan flowchart selesainya.

4. HASIL DAN PEMBAHASAN

Untuk Mengimplementasikan Teknik Data Mining dengan metode K-Means dalam *clustering* peminjaman buku pada perpustakaan UINSU. Penulis menggunakan *software* RapidMiner untuk *clustering* data peminjaman buku agar memudahkan petugas perpustakaan dalam menentukan prioritas stok buku di perpustakaan tersebut. Adapun kriteria (atribut) yang dipakai pada penelitian ini ialah:

Tabel 1. Atribut Penelitian

Atribut		Keterangan
Kode Eksemplar		Id yang melekat pada buku
Judul		Judul dari buku yang tersedia
Peminjaman		Jumlah buku yang dipinjam
Atribut	Nilai	Keterangan
Rekomendasi	<i>Cluster 0</i>	Menambah jumlah stok buku
	<i>Cluster 1</i>	Tidak Menambah jumlah stok buku

4.1. Representasi Data

Pada penelitian ini, terdapat 2154 data yang dikumpulkan dari perpustakaan UINSU. Atribut yang digunakan meliputi judul buku, jumlah buku yang dipinjam, dan waktu peminjaman. Data ini mencerminkan pola dan kebiasaan peminjam buku di perpustakaan tersebut. Dengan menggunakan algoritma K-Means, data ini akan di-*Cluster* untuk mengidentifikasi kelompok peminjam buku yang serupa. Judul skripsi " Rekomendasi Penambahan Koleksi Buku Menggunakan Algoritma K-Means *Clustering* Di Perpustakaan UINSU " menandakan tujuan penelitian untuk memberikan rekomendasi penambahan koleksi buku berdasarkan pola peminjaman yang ditemukan melalui analisis *cluster*. Ini dapat menjadi

pendekatan yang efektif untuk meningkatkan kepuasan pengguna dan keefektifan koleksi buku di perpustakaan UINSU.

Tabel 2. Data Set

No	ID	Jumlah Buku Dipinjam	Waktu Peminjaman
1	284865	1	7
2	211030	1	7
3	285236	7	49
4	284793	2	14
5	228992	2	14
6	200868	2	14
7	206036	1	7
8	205510	2	11
9	228788	2	12
10	207077	4	28
11	253277	4	28
12	206496	1	7
13	270971	1	5
14	256255	6	42
15	238868	1	7

4.2. Implementasi K-Means

Setelah menentukan dataset, maka perlu menentukan jumlah *cluster* yang akan dibentuk. Adapun *cluster* yang akan dibentuk antara lain:

- Cluster* 1 (C0) = Tidak Rekomendasi
- Cluster* 2 (C1) = Rekomendasi

Menetapkan C pusat centroid awal secara *random* dari dataset diatas terpilih 2 *cluster* dan pusat centroid diantaranya :

Tabel 3. Centroid Awal

<i>Cluster</i> 0	1	7
<i>Cluster</i> 1	7	49

Mengalokasikan semua data/obyek ke dalam *cluster* terdekat. Berikut hasil dari alokasi data ke jarak *cluster*. Adapun hasil dari jarak ke *cluster* diperoleh dari perhitungan dengan rumus :

$$d_{x,y} = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$$

$$d(1,0) = \sqrt{(1-1)^2 + (7-7)^2} = 0$$

$$d(1,1) = \sqrt{(1-7)^2 + (7-49)^2} = 42.42640687$$

$$d(2,0) = \sqrt{(1-1)^2 + (7-7)^2} = 0$$

$$d(2,1) = \sqrt{(1-7)^2 + (7-49)^2} = 42.42640687$$

$$d(3,0) = \sqrt{(7-1)^2 + (49-7)^2} = 42.42640687$$

$$d(3,1) = \sqrt{(7-7)^2 + (49-49)^2} = 0$$

$$d(4,0) = \sqrt{(2-1)^2 + (14-7)^2} = 7.071067812$$

$$d(4,1) = \sqrt{(2-7)^2 + (14-49)^2} = 35.35533906$$

$$d(5,0) = \sqrt{(2-1)^2 + (14-7)^2} = 7.071067812$$

$$d(5,1) = \sqrt{(2-7)^2 + (14-49)^2} = 35.35533906$$

$$d(6,0) = \sqrt{(2-1)^2 + (14-7)^2} = 7.071067812$$

$$d(6,1) = \sqrt{(2-7)^2 + (14-49)^2} = 35.35533906$$

$$d(7,0) = \sqrt{(1-1)^2 + (7-7)^2} = 0$$

$$d(7,1) = \sqrt{(1-7)^2 + (7-49)^2} = 42.42640687$$

$$d(8,0) = \sqrt{(2-1)^2 + (11-7)^2} = 4.123105626$$

$$d(8,1) = \sqrt{(2-7)^2 + (11-49)^2} = 38.327535795$$

$$d(9,0) = \sqrt{(2-1)^2 + (12-7)^2} = 5.099019514$$

$$d(9,1) = \sqrt{(2-7)^2 + (12-49)^2} = 37.33630941$$

$$d(10,0) = \sqrt{(4-1)^2 + (28-7)^2} = 21.21320344$$

$$d(10,1) = \sqrt{(4-7)^2 + (28-49)^2} = 21.21320344$$

$$d(11,0) = \sqrt{(4-1)^2 + (28-7)^2} = 21.21320344$$

$$d(11,1) = \sqrt{(4-7)^2 + (28-49)^2} = 21.21320344$$

$$d(12,0) = \sqrt{(1-1)^2 + (7-7)^2} = 0$$

$$d(12,1) = \sqrt{(1-7)^2 + (7-49)^2} = 42.42640687$$

$$d(13,0) = \sqrt{(1-1)^2 + (5-7)^2} = 2$$

$$d(13,1) = \sqrt{(1-7)^2 + (5-49)^2} = 44.40720662$$

$$d(14,0) = \sqrt{(6-1)^2 + (42-7)^2} = 35.35533906$$

$$d(14,1) = \sqrt{(6-7)^2 + (42-49)^2} = 7.071067812$$

$$d(15,0) = \sqrt{(1-1)^2 + (7-7)^2} = 0$$

$$d(15,1) = \sqrt{(1-7)^2 + (7-49)^2} = 42.42640687$$

Setelah melakukan perhitungan maka didapat hasil seperti berikut ini :

Tabel 4. Hasil Jarak ke Cluster

No	ID	Jumlah Buku Dipinjam	Waktu Peminjaman	Jarak ke Cluster		Hasil Cluster
				C0	C1	
1	284865	1	7	0	42.42641	0
2	211030	1	7	0	42.42641	0
3	285236	7	49	42.42641	0	1
4	284793	2	14	7.071068	35.35534	0
5	228992	2	14	7.071068	35.35534	0
6	200868	2	14	7.071068	35.35534	0
7	206036	1	7	0	42.42641	0
8	205510	2	11	4.123106	38.32754	0
9	228788	2	12	5.09902	37.33631	0
10	207077	4	28	21.2132	21.2132	0
11	253277	4	28	21.2132	21.2132	0
12	206496	1	7	0	42.42641	0
13	270971	1	5	2	44.40721	0
14	256255	6	42	35.35534	7.071068	1
15	238868	1	7	0	42.42641	0

Menentukan kembali titik pusat *centroid* yang baru berdasarkan rata-rata. Nilai rata-rata yang dimaksud ialah titik awal *centroid* yang di dapat kemudian *centroid* baru tersebut didapat dari rumus = nilai hasil /banyak hasil.

Centroid baru tersebut didapat dari rumus = nilai hasil /banyak hasil.

$$\begin{aligned} \text{Cluster 0 (Jumlah Buku Dipinjam)} &= (1+1+2+2+2+1+2+2 +4+4+1+1+1)/13 \\ &= 1.84615385 \end{aligned}$$

$$\begin{aligned} \text{Cluster 0 (Waktu Peminjaman)} &= (7+7+14+14+14+7+11+12 +28+28+7+5+7)/13 \\ &= 12.3846154 \end{aligned}$$

$$\text{Cluster 1 (Jumlah Buku Dipinjam)} = (7+6)/2= 6.5$$

$$\text{Cluster 1 (Waktu Peminjaman)} = (49+42)/2= 45,5$$

Sehingga didapat nilai centroid baru antara lain :

Tabel 5. Centroid Baru

Cluster 0	1.84615385	12.3846154
Cluster 1	6.5	45,5

Melakukan perhitungan jarak ke *cluster* seperti yang dilakukan pada proses awal:

$$d(1,0) = \sqrt{(1 - 1.84615385)^2 + (7 - 12.3846154)^2} = 5.450693474$$

$$d(1,1) = \sqrt{(1 - 6.5)^2 + (7 - 45,5)^2} = 38.89087297$$

$$d(2,0) = \sqrt{(1 - 1.84615385)^2 + (7 - 12.3846154)^2} = 5.450693474$$

$$d(2,1) = \sqrt{(1 - 6.5)^2 + (7 - 45,5)^2} = 38.89087297$$

$$d(3,0) = \sqrt{(7 - 1.84615385)^2 + (49 - 12.3846154)^2} = 36.97632377$$

$$d(3,1) = \sqrt{(7 - 6.5)^2 + (49 - 45.5)^2} = 3.535533906$$

$$d(4,0) = \sqrt{(2 - 1.84615385)^2 + (14 - 12.3846154)^2} = 1.62269407$$

$$d(4,1) = \sqrt{(2 - 6.5)^2 + (14 - 45.5)^2} = 31.81980515$$

$$d(5,0) = \sqrt{(2 - 1.84615385)^2 + (14 - 12.3846154)^2} = 1.62269407$$

$$d(5,1) = \sqrt{(2 - 6.5)^2 + (14 - 45.5)^2} = 31.81980515$$

$$d(6,0) = \sqrt{(2 - 1.84615385)^2 + (14 - 12.3846154)^2} = 1.62269407$$

$$d(6,1) = \sqrt{(2 - 6.5)^2 + (14 - 45.5)^2} = 31.81980515$$

$$d(7,0) = \sqrt{(1 - 1.84615385)^2 + (7 - 12.3846154)^2} = 5.450693474$$

$$d(7,1) = \sqrt{(1 - 6.5)^2 + (7 - 45.5)^2} = 38.89087297$$

$$d(8,0) = \sqrt{(2 - 1.84615385)^2 + (11 - 12.3846154)^2} = 1.39313619$$

$$d(8,1) = \sqrt{(2 - 6.5)^2 + (11 - 45.5)^2} = 34.79224051$$

$$d(9,0) = \sqrt{(2 - 1.84615385)^2 + (12 - 12.3846154)^2} = 0.41424346$$

$$d(9,1) = \sqrt{(2 - 6.5)^2 + (12 - 45.5)^2} = 33.80088756$$

$$d(10,0) = \sqrt{(4 - 1.84615385)^2 + (28 - 12.3846154)^2} = 15.76322586$$

$$d(10,1) = \sqrt{(4 - 6.5)^2 + (28 - 45.5)^2} = 17.67766953$$

$$d(11,0) = \sqrt{(4 - 1.84615385)^2 + (28 - 12.3846154)^2} = 15.76322586$$

$$d(11,1) = \sqrt{(4 - 6.5)^2 + (28 - 45.5)^2} = 17.67766953$$

$$d(12,0) = \sqrt{(1 - 1.84615385)^2 + (7 - 12.3846154)^2} = 5.450693474$$

$$d(12,1) = \sqrt{(1 - 6.5)^2 + (7 - 45.5)^2} = 38.89087297$$

$$d(13,0) = \sqrt{(1 - 1.84615385)^2 + (5 - 12.3846154)^2} = 7.432934881$$

$$d(13,2) = \sqrt{(1 - 6.5)^2 + (5 - 45.5)^2} = 40.87175064$$

$$d(14,0) = \sqrt{(6 - 1.84615385)^2 + (42 - 12.3846154)^2} = 29.90527$$

$$d(14,1) = \sqrt{(6 - 6.5)^2 + (42 - 45.5)^2} = 3.535533906$$

$$d(15,0) = \sqrt{(1 - 1.84615385)^2 + (7 - 12.3846154)^2} = 5.450693474$$

$$d(15,1) = \sqrt{(1 - 6.5)^2 + (7 - 45.5)^2} = 38.89087297$$

setelah melakukan perhitungan maka didapat hasil seperti berikut ini:

Tabel 6. Hasil Jarak ke Cluster

No	ID	Jumlah Buku Dipinjam	Waktu Peminjaman	Jarak ke Cluster		Hasil Cluster
				C0	C1	
1	284865	1	7	5.450693	38.89087	0
2	211030	1	7	5.450693	38.89087	0
3	285236	7	49	36.97632	3.535534	1
4	284793	2	14	1.622694	31.81981	0
5	228992	2	14	1.622694	31.81981	0
6	200868	2	14	1.622694	31.81981	0
7	206036	1	7	5.450693	38.89087	0
8	205510	2	11	1.393136	34.79224	0
9	228788	2	12	0.414243	33.80089	0
10	207077	4	28	15.76323	17.67767	0
11	253277	4	28	15.76323	17.67767	0
12	206496	1	7	5.450693	38.89087	0
13	270971	1	5	7.432935	40.87175	0
14	256255	6	42	29.90527	3.535534	1
15	238868	1	7	5.450693	38.89087	0

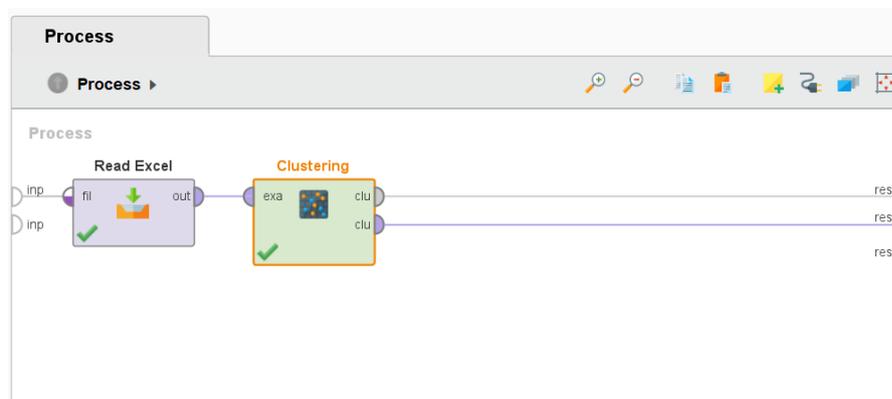
Dalam perhitungan K-means, jika *centroid* belum sesuai dengan titik awal maka proses perhitungan akan berlanjut dan jika titik *centroid* sudah sama dengan titik awal maka proses perhitungan berhenti. Hasil dari tahapan yang pertama dan kedua tidak memiliki perubahan *centroid* maka perhitungan selesai dengan hasil yang didapat.

4.3. Penerapan RapidMiner

Tahapan yang akan dilakukan selanjutnya adalah penerapan. Penerapan bertujuan untuk melihat seberapa jauh sistem yang telah dibangun dengan sistem yang diharapkan. Pada penelitian ini peneliti menggunakan *RapidMiner* dalam penerapan metode K-Means.

1. Tampilan Operator RapidMiner

Operator RapidMiner pada gambar berfungsi untuk menciptakan pusat data pada dataset penelitian yang berasal dari file CSV.

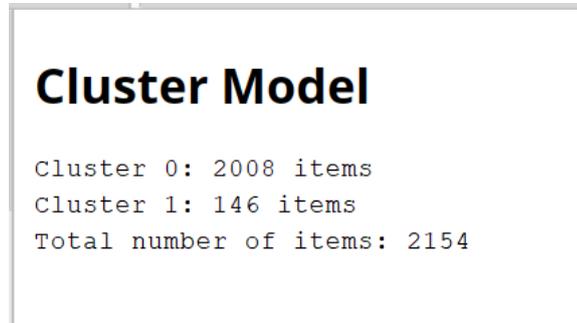


Gambar 3. Tampilan Operator RapidMiner

Pada gambar diatas merupakan operator proses *RapidMiner* yaitu komponen utama dalam merancang alur kerja analisis data. Setelah memasukkan data Excel dengan informasi judul buku, banyak buku dipinjam, dan waktu peminjaman, operator k-means dapat digunakan

untuk mengelompokkan data menjadi dua *Cluster* berdasarkan pola yang ada. Operator ini membantu mengidentifikasi struktur internal dalam dataset dan memudahkan pemahaman tentang bagaimana buku-buku tersebut dapat dikelompokkan berdasarkan pola peminjaman dan waktu peminjaman.

2. Tampilan Hasil Pengelompokan K-Means



Gambar 4. Tampilan Hasil Pengelompokan K-Means

Hasil dari *RapidMiner Cluster Model* menunjukkan bahwa data buku telah berhasil dikelompokkan menjadi dua *cluster*. *Cluster 0* memiliki 2008 buku, sementara *Cluster 1* memiliki 146 buku, dengan total keseluruhan 2154 buku. Pengelompokan ini didasarkan pada atribut yang digunakan, yaitu judul buku, jumlah buku yang dipinjam, dan waktu peminjaman. Dengan demikian, kita dapat menyimpulkan bahwa ada dua pola atau karakteristik utama dalam perilaku peminjaman buku yang tercermin dalam *cluster* tersebut.

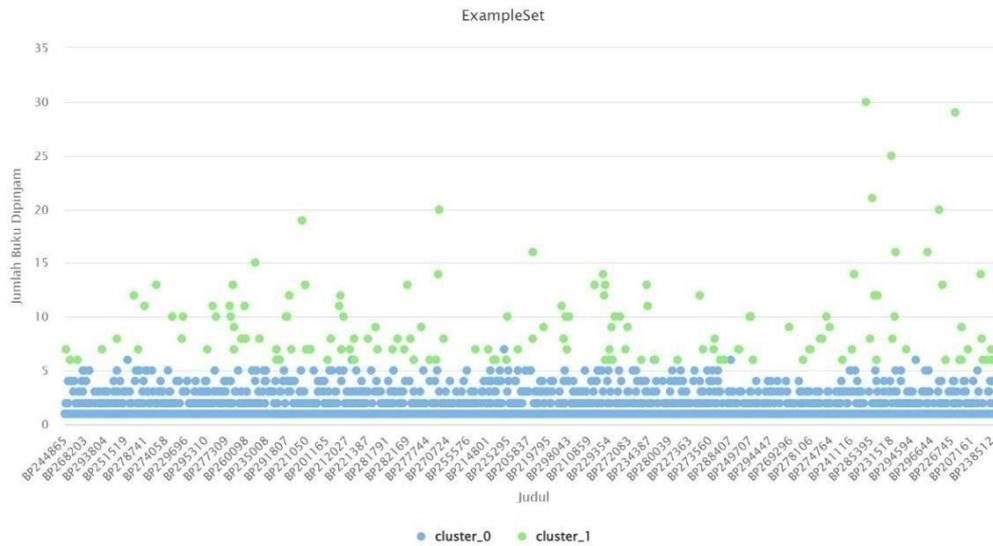
3. Tampilan Tabel Centroid RapidMiner

Attribute	cluster_0	cluster_1
Judul	1060.222	1110.671
Jumlah Buku Dipinjam	1.733	9.397
Waktu Peminjaman	11.562	63.562

Gambar 5. Tampilan Tabel Centroid RapidMiner

Dalam hasil *RapidMiner*, tabel centroid menampilkan nilai-nilai atribut untuk masing-masing *cluster*. Atribut termasuk Judul, Jumlah Buku Dipinjam, dan Waktu Peminjaman. Untuk setiap atribut, terdapat nilai centroid yang mewakili pusat *cluster* tersebut. Sebagai contoh, dalam *Cluster 0*, nilai centroid untuk Judul, Jumlah Buku Dipinjam, dan Waktu Peminjaman berturut-turut adalah 1060.22, 1.73, dan 11.56. Sedangkan dalam *Cluster 1*, nilai centroid yang sesuai adalah 1110.67, 9.40, dan 63.56. Informasi ini memberikan gambaran tentang karakteristik rata-rata atau pusat massa dari data dalam masing-masing *cluster*, mempermudah interpretasi hasil analisis *cluster* tersebut.

4. Tampilan Visualizations RapidMiner



Gambar 6. Tampilan Visualizations RapidMiner

Pada gambar Visualisasi dalam analisis *cluster* memegang peran kunci dengan berbagai kegunaan, termasuk membantu pemahaman struktur data, penentuan jumlah *cluster* yang optimal melalui teknik seperti "*Elbow*," evaluasi kualitas *cluster*, pemilihan fitur penting, identifikasi outlier, dan presentasi hasil dengan cara yang mudah dipahami. Dengan memanfaatkan grafik dan *plot*, visualisasi memungkinkan pengguna untuk melihat pola, hubungan, dan karakteristik *cluster* dengan lebih jelas, serta memberikan pemahaman yang mendalam tentang data yang dianalisis.

5. Hasil

Tahapan ini berisi mengenai hasil yang telah diperoleh peneliti dari pengujian terhadap penerapan algoritma K-Means menggunakan aplikasi RapidMiner dalam Rekomendasi Penambahan Koleksi Buku Menggunakan Algoritma K-Means *Clustering* Di Perpustakaan UINSU.

Tabel 7. Hasil Akhir

No	ID	Jumlah Buku Dipinjam	Waktu Peminjaman	Hasil <i>Cluster</i>
1	284865	1	7	Tidak Rekomendasi
2	211030	1	7	Tidak Rekomendasi
3	285236	7	49	Rekomendasi
4	284793	2	14	Tidak Rekomendasi
5	228992	2	14	Tidak Rekomendasi
6	200868	2	14	Tidak Rekomendasi
7	206036	1	7	Tidak Rekomendasi
8	205510	2	11	Tidak Rekomendasi
9	228788	2	12	Tidak Rekomendasi
10	207077	4	28	Tidak Rekomendasi
11	253277	4	28	Tidak Rekomendasi

12	206496	1	7	Tidak Rekomendasi
13	270971	1	5	Tidak Rekomendasi
14	256255	6	42	Rekomendasi
15	238868	1	7	Tidak Rekomendasi

5. KESIMPULAN

Penerapan algoritma K-Means telah berhasil diimplementasikan pada dataset dalam Rekomendasi Penambahan Koleksi Buku Di Perpustakaan UINSU. Berdasarkan hasil RapidMiner *cluster* Model, dapat disimpulkan bahwa data peminjaman buku dari perpustakaan UINSU telah berhasil dikelompokkan menjadi dua *cluster* yaitu *Cluster 0* (Tidak Rekomendasi) dan *Cluster 1* (Rekomendasi). *Cluster 0* (Tidak Rekomendasi) memiliki 2008 item, sementara *Cluster 1* (Rekomendasi) memiliki 146 item, dengan total 2154 item keseluruhan. Analisis Centroid menunjukkan karakteristik rata-rata atau pusat massa dari masing-masing *cluster*. Sebagai contoh, *Cluster 0* memiliki nilai Centroid untuk Judul, Jumlah Buku Dipinjam, dan Waktu Peminjaman masing-masing adalah 1060.22, 1.73, dan 11.56. Di sisi lain, *Cluster 1* memiliki nilai Centroid yang berbeda, yaitu 1110.67, 9.40, dan 63.56. Hasil ini memberikan pemahaman yang lebih baik tentang pola peminjaman buku dan memungkinkan perpustakaan UINSU untuk membuat rekomendasi penambahan koleksi buku yang lebih tepat dan sesuai dengan preferensi peminjam dalam masing-masing *Cluster*.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak para peninjau yang telah berkontribusi dalam proses peer review naskah-naskah dalam hal ini masalah. Dukungan profesional dan bantuan dari semua reviewer yang dihormati telah membuat jurnal ini memenuhi syarat untuk diterbitkan. Serta dosen dan teman yang juga memberikan dukungan pada penulisan penelitian ini.

REFERENSI

- [1] Bakker, E. (2020). Implementasi Data Mining Clustering Data Perpustakaan Menggunakan Algoritma K-Means untuk Menentukan Penambahan Koleksi Buku di Perpustakaan UPY. *SEMINAR NASIONAL Dinamika Informatika*, 22–25.
- [2] Firmansyah, T., Poningsih, P., & ... (2022). Analisis Clustering Algoritma K-Means Sebagai Rekomendasi Penambahan Koleksi Buku Di Perpustakaan Madrasah Tsanawiyah Negeri 2 Simalungun. *ZAHRA: Buletin Big Data ...*, 1(1), 44–48. <https://ejurnal.pdsi.or.id/index.php/zahra/article/view/13%0Ahttps://ejurnal.pdsi.or.id/index.php/zahra/article/viewFile/13/11>
- [3] Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 5(1), 17–24. <https://doi.org/10.25077/teknosi.v5i1.2019.17-24>
- [4] Harahap, B. (2019). Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung). *Regional Development Industry & Health Science, Technology and Art of Life*, 394–403. <https://ptki.ac.id/jurnal/index.php/readystar/article/view/82>

- [5] Mustika, R. D., Zakir, A., Informasi, S., Harapan, U., & Skripsi, J. (2022). *JURNAL MEDIA INFORMATIKA [JUMIN] Implementasi Algoritma K-Means Untuk Clustering Judul Skripsi Universitas Harapan Medan JURNAL MEDIA INFORMATIKA [JUMIN]*. 4, 40–47.
- [6] Nanda, A. P., Pramono, D. E. H., & Hartati, S. (2020). Menentukan Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Metode Algoritma K-Means. *Jurnal Sistem Informasi Dan Telematika*, 11(1), 23–28.
- [7] Nasir, J. (2021). Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 11(2), 690–703. <https://doi.org/10.24176/simet.v11i2.5482>
- [8] Prastiwi, H., Pricilia, J., & Raswir, E. (2022). Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM). *Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)*, 1(April), 141–148.
- [9] Rahma, L. M. (2020). *Implementasi Algoritma K-Means Berbasis Android Untuk Clustering Tingkat Pemahaman Mahasiswa Skripsi Oleh : Lutfia Miftahur Rahma*.
- [10] Ramadanti, E., & Muslih, M. (2022). Penerapan Data Mining Algoritma K-Means Clustering Pada Populasi Ayam Petelur Di Indonesia. *Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab*, 7(1), 1–7. <https://doi.org/10.36341/rabit.v7i1.2155>
- [11] Rismayati, R., Ismail Marzuki, J., & Tapen, K. (2021). Implementasi Algoritma Frequent Pattern-Growth Terhadap Pola Mahasiswa Lulusan Dengan Rapidminer. *Jurnal Informatika & Rekayasa Elektronika*, 4(2), 106–114. <http://e-journal.stmiklombok.ac.id/index.php/jireISSN.2620-6900>
- [12] Rosadi, M., Aulia Nurhasanah, D., & Siddik Hasibuan, M. (2023). Clustering Panjang Ruas Jalan di BBPJN Sumut Menggunakan Algoritma K-Means. *Journal of Computer Science and Informatics Engineering (CoSIE)*, 02(1), 29–38. <https://doi.org/10.55537/cosie.v2i1.567>
- [13] Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner. *JBASE - Journal of Business and Audit Information Systems*, 4(1), 13–21. <https://doi.org/10.30813/jbase.v4i1.2729>
- [14] Triyansyah, D., & Fitriah, D. (2018). Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing. *Jurnal Telekomunikasi Dan Komputer*, 8(3), 163. <https://doi.org/10.22441/incomtech.v8i3.4174>
- [15] Yulia, & Silalahi, M. (2021). Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means. *Indonesian Journal of Computer Science*, 10(1). <https://doi.org/10.33022/ijcs.v10i1.3008>