

Penerapan Metode K-Means Dalam Pengelompokkan Buku Untuk Menentukan Minat Baca Pada Perpustakaan Daerah Kota Medan

Misael Oktavianda Harefa, Soeb Aripin*

Fakultas Ilmu Komputer dan Teknologi Informasi, Prodi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia
Jl. Sisingamangaraja No. 338, Medan, Sumatera Utara, Indonesia
Email: ¹misael031099@gmail.com, ^{2,*}soebaripin@gmail.com
Email korespondensi: soebaripin@gmail.com

Abstrak—Perpustakaan merupakan suatu sarana yang berfungsi sebagai pusat informasi, sumber dari berbagai ilmu pengetahuan, penelitian, rekreasi dan pelestarian budaya. Perpustakaan Daerah Kota Medan merupakan perpustakaan yang diselenggarakan oleh pemerintah daerah dengan menggunakan dana umum yang bertujuan untuk melayani masyarakat mendapatkan informasi secara menyeluruh tanpa membedakan jenis kelamin, agama, ras, usia, pekerjaan dan kedudukan. Pada tahun 2021 perpustakaan daerah kota medan memiliki 28 ribuan judul buku dalam beberapa kategori. Dari banyaknya buku yang terdapat pada perpustakaan tersebut maka harus dibutuhkan sebuah system yang dimana system tersebut bermanfaat bagi pihak perpustakaan maupun si pembaca dalam memaksimalkan melakukan pengelompokkan buku maupun pencarian buku dengan mudah oleh si pembaca, maka dari itu digunakan metode K-Means Clustering, dimana metode ini merupakan salah satu metode dalam data mining yang melakukan proses data clustering yang dikelompokkan dalam satu bahkan lebih cluster. Pada penelitian ini menggunakan sebanyak 100 sampel data kategori buku yang ada pada perpustakaan Kota Medan. Penelitian ini mengelompokkan data kategori menjadi 3 cluster yaitu paling diminati, diminati dan kurang diminati. Hasil dari proses metode ini akan mengetahui data kategori buku yang paling diminati sehingga kedepannya akan menjadi pertimbangan pihak pustakawan untuk meningkatkan koleksi buku pada perpustakaan daerah Kota Medan. Proses perhitungan metode K-Means dalam pengelompokkan buku hanya dilakukan sampai iterasi ke-2 dikarenakan iterasi -3 mendapatkan hasil nilai yang sama. Pada Kluster 1 (Paling Diminati) memilih 6 kategori buku diantaranya Kategori 020-Perpustakaan dan Informasi, 070-Media Massa, Jurnalisme dan Publikasi, 050-Psikologi, 420-Bahasa Indonesia, 600-Teknologi, 650-Manajemen. Pada Kluster 2 (Diminati) memilih 16 kategori buku diantaranya 000-Publikasi Umum dan Informasi Umum, 030-Ensiklopedia dan Buku, 040-Biografi, 050-Majalah dan Jurnal, 090-Manuskrip dan Buku Langka, 210-Agama Islam, 300-Ilmu Sosial, 320-Ilmu Politik, 330-Ilmu Ekonomi, 410-Bahasa Indonesia, 510-Matematika, 620-Ilmu Teknik, 770-Fotografi dan Foto, 780-Musik, 910-Geografi Umum Perjalanan, 930-Sejarah Dunia Lama. Dan pada kluster 3 (Kurang Diminati) terpilih sebanyak 78 kategori buku. 10 diantaranya 010-Bibliografi, 060-Asosiasi Organisasi dan Museum, 080-Kutipan, 100-Filsafat dan Psikologi, 110-Matematika, 120-Epistimologi, 030-Parapsikologi dan Okultisme, 040-Pemikiran Filosofis, 060-Filosofi Logis, 070-Etik.

Kata Kunci: Metode K-Means Pengelompokkan Buku, Minat Baca, Perpustakaan

Abstract—The library is a facility that functions as an information center, a source of various knowledge, research, recreation and cultural preservation. The Medan City Regional Library is a library organized by the regional government using general funds which aims to serve the public in obtaining comprehensive information without distinguishing gender, religion, race, age, occupation and position. In 2021 the Medan city regional library has 28 thousands of titles books in several categories. Of the many books contained in the library, a system must be needed where the system is useful for both the library and the reader in maximizing grouping of books and searching books easily by the reader, therefore the K-Means Clustering method is used, where the method This is a method in data mining that processes clustering data that is grouped into one or more clusters. In this study, 100 samples of book category data were used in the Medan City library. This study groups the data categories into 3 clusters, namely the most desirable, desirable and least desirable. The results of this method process will find out the most popular book category data so that in the future it will be a consideration for the librarian to increase the collection of books at the Medan City regional library. The process of calculating the K-Means method in grouping books is only carried out until the 2nd iteration because iteration -3 gets the same value. In Cluster 1 (Most Interested) choose 6 categories of books including Category 020-Library and Information, 070-Mass Media, Journalism and Publication, 050-Psychology, 420-Indonesian, 600-Technology, 650-Management. In Cluster 2 (Desired) choose 16 categories of books including 000-General Publications and General Information, 030-Encyclopedias and Books, 040-Biography, 050-Magazines and Journals, 090-Manuscripts and Rare Books, 210-Islamic Religion, 300-Science Social, 320-Political Science, 330-Economics, 410-Indonesian, 510-Mathematics, 620-Technical Sciences, 770-Photography and Photos, 780-Music, 910-General Travel Geography, 930-Old World History. And in cluster 3 (Less Interested) 78 categories of books were selected. 10 of them 010-Bibliography, 060-Association of Organizations and Museums, 080-Quotes, 100-Philosophy and Psychology, 110-Mathematics, 120-Epistimology, 030-Parapsychology and Occultism, 040-Philosophical Thought, 060 Logical Philosophy, 070-Ethics.

Keywords: K-Means, Method of Grouping Books, Interest in Reading, Libraries

1. PENDAHULUAN

Perkembangan teknologi beriring sejalan dengan perkembangan ilmu pengetahuan. Dengan perkembangan teknologi zaman sekarang ini sangat membuka banyak peluang berinovasi tentunya dalam berbagai bidang keilmuan salah satunya adalah ilmu data mining. Mulai dari Analisis, Clustering hingga prediksi berdasarkan parameter yang telah ditentukan [1].

Perpustakaan Daerah Kota Medan merupakan perpustakaan yang diselenggarakan oleh pemerintah daerah dengan menggunakan dana umum yang bertujuan untuk melayani masyarakat mendapatkan informasi secara menyeluruh tanpa membedakan jenis kelamin, agama, ras, usia, pekerjaan dan kedudukan. Tujuan perpustakaan daerah

memberikan kesempatan kepada warga masyarakat untuk menggunakan bahan Pustaka dalam meningkatkan pengetahuan, keterampilan dan kesejahteraannya. Salah satu untuk mencapai tujuan dan fungsi suatu perpustakaan umum adalah ketersediaan bahan Pustaka yang dibutuhkan oleh pemustak [2].

Dalam Pemenuhan ketersediaan bahan Pustaka pengelolaanya masih dianggap belum optimal karena terjadi ketidakseimbangan beberapa buku antara rasio peminjam dengan stok buku yang ada sehingga terjadi kekosongan apabila jumlah eksemplar tidak diperbanyak. Disisi lain buku-buku yang mengalami penambahan stok judul buku tidak sinkron dengan apa yang diinginkan si peminat baca untuk mendapatkan buk tersebut. Terkait hal ini maka dibutuhkan data mining untuk menyelesaikan permasalahan tersebut. Metode yang digunakan adalah metode K-Means [3].

Metode K-Means merupakan salah satu metode clustering yang digunakan untuk mengklusterataupun mengelompokkan data menjadi beberapa data. Metode ini mengelompokkan jumlah data yang besar berdasarkan titik centroid (Pusat). Data yang akan dicluster diambil secara random. Dengan penerapan metode K-Means diharapkan dapat membantu mengelompokkan buku yang paling diminati maupun tidak diminati pada perpustakaan daerah kota medan, sehingga pihak pengelola dapat dengan mudah melakukan analisis dan mengambil keputusan [4].

Berdasarkan Penelitian yang pernah dilakukan sebelumnya oleh Nisriina Nur Hasanah dan Agus Siddiq Purnomo yang berjudul “Implementasi Data Mining Untuk Pengelompokkan Buku Menggunakan Algoritma K-Means Clustering (Studi Kasus : Perpustakaan Politeknik LPP Yogyakarta)” mendapatkan solusi dari hasil penelitian bahwa dari 30 judul buku yang diuji, 3 diantaranya merupakan judul buku yang paling diminati seperti Buku Pintar Mandor Seri Budidaya Tanaman Tebu, dan Buku Pintar Mandor Seri Budidaya Tanaman Kelapa Sawit dan Kupas Tuntas Agribisnis Kelapa Sawit[5].

Selanjutnya pada Penelitian Yulia Darmi dan Agus Setiawan pada tahun 2016 tentang “Penerapan Metode Clustering K-Means Dalam Pengelompokkan Penjualan Produk” menarik kesimpulan bahwa pengklasteran dengan menggunakan metode K-Means didapatkan untuk produk yang laku yaitu terdiri dari makanan dan minuman dengan yang tidak laku adalah produk kosmetik [4]

Dan Pada penelitian ilmiah terakhir yang dilakukan oleh Muhammad Rafi Muttaqin dan Meriska Defriani dengan penelitian yang berjudul “Algoritma K-Means Untuk Pengelompokkan Topik Skripsi Mahasiswa” berkesimpulan dengan mengelompokkan nilai clusteri paling tinggi dapat menunjukkan kemampuan mahasiswa pada tiap kelompok bidang keahlian. Hasil pengelompokkan ini dilihat dengan melihat jumlah mata kuliah pada setiap kelompok bidang keahlian[6].

Berdasarkan permasalahan dan penelitian terkait diatas tujuan penelitian ini untuk pengelompokan minat pembaca buku di perpustakaan daerah kota medan dengan menerapkan metode K-mens sehingga dapat memberikan informasi yang bermanfaat bagi pihak terkait.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Secara sederhana Data Mining adalah proses untuk menambang atau menggali informasi yang tersembunyi dari bongkahan data besar. Inti dari proses proses Knowledge Discovery in Database (KDD) adalah Data mining, yang algoritmanya mengeksplor dan membangun model data. Knowledge Discovery in Database (KDD) sebagai pengorganisasian untuk proses pengidentifikasian data yang benar dan berguna untuk penemuan pola dari kumpulan data yang besar dan kompleks[7][8].

Data mining adalah proses pencarian pola atau informasi yang menarik didalam suatu data dengan menggunakan teknik atau metode tertentu. Dalam data mining Teknik, metode, atau algoritma itu bervariasi. Pemilihan metode atau algoritma yang tepat bergantung pada tujuan dan proses Knowledge Discovery in Database (KDD) secara keseluruhan[9][10].

2.2 Pengelompokkan Buku

Klasifikasi berasal dari kata Latin “*classis*” yang artinya adalah proses pengelompokan. Dalam bahasa Inggris yaitu classification yang artinya adalah penggolongan atau pembagian. Dalam bahasa Indonesia klasifikasi berarti pengelompokan menurut jenis. Menurut Sulisty-Basuki dalam bukunya yang berjudul “*Pengantar Ilmu Perpustakaan*” klasifikasi adalah proses pengelompokan, artinya mengumpulkan benda yang sama serta memisahkan benda yang tidak sama. Klasifikasi pada dunia perpustakaan dan informasi didefinisikan sebagai: penyusunan sistematis terhadap buku dan bahan pustaka lain atau katalog/entri indeks berdasarkan subyek, dalam cara paling berguna bagi mereka yang membaca atau mencari informasi[11][1].

2.3 Minat Baca

Minat baca merupakan suatu ketertarikan untuk dapat mengartikan atau menafsirkan media kata-kata dengan tujuan untuk memperoleh informasi yang dibutuhkan. Dengan adanya minat baca dapat mendorong seseorang untuk giat memperluas pengetahuannya[12].

2.4 Metode K-Means Clustering

Clustering merupakan salah satu algoritma dalam data mining yang bertujuan sebagai pengelompokan beberapa objek ke dalam cluster-cluster yang karakteristiknya sama, mengelompokkan item data ke dalam sejumlah kecil cluster. Cluster yang sama dan disesuaikan terhadap objek-objek yang berbeda cluster].

Algoritma K-Means data diklasifikasikan berdasarkan kedekatannya dengan centroid. Data dikelompokkan dengan memaksimalkan data yang sama dalam satu klaster dan meminimalkan data yang sama antar klaster. Ukuran kesamaan data yang digunakan dalam klaster merupakan fungsi jarak.

Suatu metode yang menganalisa data atau metode dalam data mining yang melakukan proses data clustering non-hirarki yang dimana data dikelompokkan dalam satu bahkan lebih cluster dan merupakan salah satu metode yang dilakukan dengan sistem partisi adalah K-Means Clustering [13][14][15]. Langkah-langkah melakukan clustering dengan metode K-Means adalah:

- Tentukan jumlah nilai k sebagai jumlah cluster.
- Alokasikan data kedalam kelompok secara random.
- Hitung pusat cluster (centroid) menggunakan mean untuk masing-masing cluster dengan persamaan Euclidean yaitu sebagai berikut:

$$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2} \quad (1)$$

Dimana $D(i,j)$ yaitu Jarak data ke i ke pusat cluster j, X_{ki} yaitu Data ke i pada atribut data ke k, X_{kj} yaitu Titik pusat ke j pada atribut ke k

- Alokasikan data berdasarkan jarak terdekat antara data dengan centroidnya.
- Kembali kelangkah sebelumnya, jika ternyata masih ada data yang berpindah cluster atau jika nilai centroid diatas nilai ambang, atau jika nilai pada fungsi objektif yang digunakan masih diatas ambang. Pada penerapan metode k-means cluster analysis, data yang bisa diolah dalam perhitungan adalah data numerik yang berbentuk angka. Sedangkan data selain angka juga bisa diterapkan tetapi terlebih dahulu harus dilakukan pengkodean untuk mempermudah perhitungan jarak/kesamaan karakteristik yang dimiliki dari setiap objek [16][17].

Setiap objek dihitung kedekatan jaraknya berdasarkan karakter yang dimiliki dengan pusat cluster yang sudah ditentukan sebelumnya, jarak terkecil antara objek dengan masing-masing cluster merupakan anggota cluster yang terdekat. Setelah jumlah cluster ditentukan, selanjutnya dipilih sebanyak 3 objek secara acak sesuai jumlah cluster yang dibentuk sebagai pusat cluster awal untuk dihitung jarak kedekatannya terhadap semua objek yang ada [18][19][20]

3. HASIL DAN PEMBAHASAN

Pada tahapan ini merupakan tahapan untuk mempersiapkan data yang diperlukan untuk proses penerapan data mining yang bertujuan agar data yang digunakan sesuai dengan permasalahan yang terjadi. Maka dari itu data yang digunakan dalam penelitian ini terdiri dari Kategori Buku, Stok Buku dan Jumlah Peminjaman pada Perpustakaan Daerah.

3.1 Menentukan Kode Kriteria dan Kategori Buku

Tabel 1. Kode Kriteria

Kode Kriteria	Nama Kriteria
K1	Peminjam Buku
K2	Stok Buku

Selanjutnya menentukan data dari kategori buku yang didapatkan dari perpustakaan daerah seperti pada tabel 2. Data Buku dibawah ini:

Tabel 2. Data Buku

No	Kategori Buku	Peminjam Buku	Stok Buku
		(K1)	(K2)
1	000-Publikasi umum, Informasi Umum	25	32
2	010-Bibliografi	10	28
3	020-Perpustakaan dan Informasi	83	57
4	030-Ensiklopedia dan Buku	31	17
5	040-Biografi	66	15
6	050-Majalah dan Jurnal	46	7
7	060-Asosiasi, Organisasi dan Museum	12	48
8	070-Media Massa, Jurnalisme dan Publikasi	54	57
9	080-Kutipan	6	10
10	090-Manuskrip dan Buku Langka	41	19
11	100-Filsafat dan Psikologi	10	5
12	110-Metafisika	6	5

No	Kategori Buku	Peminjam Buku	Stok Buku
		(K1)	(K2)
13	120-Epistimologi	15	12
14	030-Parapsikologi dan Okultisme	7	7
15	040-Pemikiran Filosofis	4	9
16	050-Psikologi	24	15
17	060-Filosofis Logis	12	4
18	070-Etik	3	1
19	180-Filosofi Kuno	5	2
20	190-Filosofi Barat Modern	1	2
21	200-Agama	15	18
22	210-Agama Islam	20	9
23	220-Alkitab	6	5
24	230-Teologi Kristen	2	6
25	240-Moral Kristen & Teologi Kebaktian	2	4
...
99	980-Asal-usul Amerika Selatan	10	2
100	990-Asal-usul Wilayah Lain	4	1

3.2 Penerapan Metode K-Means

Clustering merupakan salah satu algoritma dalam data mining yang bertujuan sebagai pengelompokan beberapa objek ke dalam cluster-cluster yang karakteristiknya sama, mengelompokkan item data ke dalam sejumlah kecil cluster. Cluster yang sama dan disesuaikan terhadap objek-objek yang berbeda cluster. Berikut ini adalah Langkah-langkah penyelesaiannya yaitu:

a. Iterasi ke-1

1. Menentukan pusat (centroid) awal cluster

Dalam menentukan pusat centroid awal ditentukan dengan acak dari sampel data yang ada. Pada penelitian ini ditentukan pusat awal seperti tabel 3 berikut:

Tabel 3. Titik Pusat (Centroid) Awal Cluster

Centroid	Data Buku	K1	K2
Centroid 1	Perpustakaan Informasi	83	57
Centroid 2	Majalah dan Jurnal	46	7
Centroid 3	Kutipan	6	10

2. Melakukan Perhitungan Jarak Dengan Pusat Cluster

Jarak antara buku dengan pusat cluster 1

$$D_{11} = \sqrt{(83 - 25)^2 + (57 - 32)^2}$$

$$D_{11} = 63,1585307$$

$$D_{12} = \sqrt{(83 - 10)^2 + (57 - 28)^2}$$

$$D_{12} = 78,54934755$$

$$D_{13} = \sqrt{(83 - 83)^2 + (57 - 57)^2}$$

$$D_{13} = 0$$

$$D_{14} = \sqrt{(83 - 31)^2 + (57 - 17)^2}$$

$$D_{14} = 65,60487787$$

$$D_{15} = \sqrt{(83 - 66)^2 + (57 - 15)^2}$$

$$D_{15} = 45,31004304$$

Jarak antara buku dengan pusat cluster 2

$$D_{21} = \sqrt{(46 - 25)^2 + (7 - 32)^2}$$

$$D_{21} = 32,64965543$$

$$D_{22} = \sqrt{(46 - 10)^2 + (7 - 28)^2}$$

$$D_{22} = 41,67733197$$

$$D_{23} = \sqrt{(46 - 83)^2 + (7 - 67)^2}$$

$$D_{23} = 62,20128616$$

$$D_{24} = \sqrt{(46 - 31)^2 + (7 - 17)^2}$$

$$D_{24} = 18,02$$

$$D_{25} = \sqrt{(46 - 66)^2 + (7 - 15)^2}$$

$$D_{25} = 18,02775638$$

Jarak antara buku dengan pusat cluster 3

$$D_{16} = \sqrt{(83 - 46)^2 + (57 - 7)^2}$$

$$D_{16} = 62,20128616$$

$$D_{17} = \sqrt{(83 - 12)^2 + (57 - 48)^2}$$

$$D_{17} = 71,56814934$$

$$D_{18} = \sqrt{(83 - 54)^2 + (57 - 57)^2}$$

$$D_{18} = 29$$

$$D_{19} = \sqrt{(83 - 6)^2 + (57 - 10)^2}$$

$$D_{19} = 90,21086409$$

$$D_{110} = \sqrt{(83 - 41)^2 + (57 - 19)^2}$$

$$D_{110} = 56,63920903$$

$$D_{26} = \sqrt{(46 - 46)^2 + (7 - 7)^2}$$

$$D_{26} = 0$$

$$D_{27} = \sqrt{(46 - 12)^2 + (7 - 48)^2}$$

$$D_{27} = 53,26349594$$

$$D_{28} = \sqrt{(46 - 54)^2 + (7 - 57)^2}$$

$$D_{28} = 50,6359556$$

$$D_{29} = \sqrt{(46 - 6)^2 + (7 - 10)^2}$$

$$D_{29} = 40,11234224$$

$$D_{210} = \sqrt{(46 - 41)^2 + (7 - 19)^2}$$

$$D_{210} = 13$$

$$D31 = \sqrt{(6 - 25)^2 + (10 - 32)^2}$$

$$D31 = 29,06888371$$

$$D32 = \sqrt{(6 - 10)^2 + (10 - 28)^2}$$

$$D32 = 18,43908891$$

$$D33 = \sqrt{(6 - 83)^2 + (10 - 67)^2}$$

$$D33 = 90,21086409$$

$$D34 = \sqrt{(6 - 31)^2 + (10 - 17)^2}$$

$$D34 = 25,96150997$$

$$D35 = \sqrt{(6 - 66)^2 + (10 - 15)^2}$$

$$D35 = 60,20797289$$

$$D36 = \sqrt{(6 - 46)^2 + (10 - 7)^2}$$

$$D36 = 40,11234224$$

$$D37 = \sqrt{(6 - 12)^2 + (10 - 48)^2}$$

$$D37 = 38,47076812$$

$$D38 = \sqrt{(6 - 54)^2 + (10 - 57)^2}$$

$$D38 = 67,17886572$$

$$D39 = \sqrt{(6 - 6)^2 + (10 - 10)^2}$$

$$D39 = 0$$

$$D310 = \sqrt{(6 - 41)^2 + (10 - 19)^2}$$

$$D310 = 36,138622$$

Kemudian setelah perhitungan cluster dilakukan, maka didapatkan hasil dari perhitungan jarak setiap data pada pusat cluster baru yang terlihat pada tabel 4 berikut:

Tabel 4. Hasil Perhitungan Pada Iterasi ke-1

No	Kategori Buku	Peminjam Buku	Stok Buku	Cluster			Jarak Terdekat
		(K1)	(K2)	C1	C2	c3	
1	000-Publikasi umum, Informasi Umum	25	32	63,1585307	32,64965543	29,06888371	29,06888371
2	010-Bibliografi	10	28	78,54934755	41,67733197	18,43908891	18,43908891
3	020-Perpustakaan dan Informasi	83	57	0	62,20128616	90,21086409	0
4	030-Ensiklopedia dan Buku	31	17	65,60487787	18,02775638	25,96150997	18,02775638
5	040-Bigrafi	66	15	45,31004304	21,54065923	60,20797289	21,54065923
6	050-Majalah dan Jurnal	46	7	62,20128616	0	40,11234224	0
7	060-Asosiasi, Organisasi dan Museum	12	48	71,56814934	53,26349594	38,47076812	38,47076812
8	070-Media Massa, Jurnalisme dan Publikasi	54	57	29	50,6359556	67,17886572	29
9	080-Kutipan	6	10	90,21086409	40,11234224	0	0
10	090-Manuskrip dan Buku Langka	41	19	56,63920903	13	36,138622	13
11	100-Filsafat dan Psikologi	10	5	89,62700486	36,05551275	6,403124237	6,403124237
12	110-Metafisika	6	5	92,91393867	40,04996879	5	5
13	120-Epistemologi	15	12	81,54140053	31,40063694	9,219544457	9,219544457
14	030-Parapsikologi dan Okultisme	7	7	90,97252332	39	3,16227766	3,16227766
15	040-Pemikiran Filosofis	4	9	92,43916919	42,04759208	2,236067977	2,236067977
16	050-Psikologi	47	50	36,67424164	43,01162634	57,28001397	36,67424164
17	060-Filosofis Logis	12	4	88,60022573	34,13209633	8,485281374	8,485281374
18	070-Etik	3	1	97,65244493	43,41658669	9,486832981	9,486832981
19	180-Filsafat Kuno	5	2	95,4410813	41,30375286	8,062257748	8,062257748
20	190-Filosofi Barat Modern	1	2	98,73702446	45,27692569	9,433981132	9,433981132
...
99	980-Asal-usul Amerika Selatan	10	2	91,40021882	36,34556369	8,94427191	8,94427191
100	990-Asal-usul Wilayah Lain	4	1	96,83491106	42,42640687	9,219544457	9,219544457

3. Melakukan Pengklasteran Nilai Cluster Iterasi 1

Setelah menghitung jarak data ke centroid, dilakukan clustering data. Nilai minimum adalah nilai yang sesuai dengan nilai opsi. Jika ditemukan nilai terkecil (minimum), maka dapat dimasukkan ke dalam cluster. C0 menunjukkan luas buku yang akan ditambah, C1 menunjukkan luas buku yang akan ditambah terlebih dahulu, dan C2 menunjukkan luas buku yang tidak ditambah. Hasil clustering dapat dilihat pada tabel 5 di bawah ini.

Tabel 5. Hasil Pengklasteran Nilai Cluster Iterasi 1

No	Kategori Buku	Peminjam Buku	Stok Buku	Cluster			Jarak Terdekat
		(K1)	(K2)	C1	C2	C3	
1	000-Publikasi umum, Informasi Umum	25	32			C3	29,06888371
2	010-Bibliografi	10	28			C3	18,43908891
3	020-Perpustakaan dan Informasi	83	57	C1			0
4	030-Ensiklopedia dan Buku	31	17		C2		18,02775638
5	040-Bigrafi	66	15		C2		21,54065923
6	050-Majalah dan Jurnal	46	7		C2		0
7	060-Asosiasi, Organisasi dan Museum	12	48			C3	38,47076812
8	070-Media Massa, Jurnalisme dan Publikasi	54	57	C1			29
9	080-Kutipan	6	10			C3	0
10	090-Manuskrip dan Buku Langka	41	19		C2		13
11	100-Filsafat dan Psikologi	10	5			C3	6,403124237
12	110-Metafisika	6	5			C3	5

No	Kategori Buku	Peminjam Buku	Stok Buku	Cluster			Jarak Terdekat
		(K1)	(K2)	C1	C2	C3	
13	120-Epistimologi	15	12			C3	9,219544457
14	030-Parapsikologi dan Okultisme	7	7			C3	3,16227766
15	040-Pemikiran Filosofis	4	9			C3	2,236067977
16	050-Psikologi	47	50	C1			36,67424164
17	060-Filosofis Logis	12	4			C3	8,485281374
18	070-Etik	3	1			C3	9,486832981
19	180-Filosofi Kuno	5	2			C3	8,062257748
20	190-Filosofi Barat Modern	1	2			C3	9,433981132
...
99	980-Asal-usul Amerika Selatan	10	2			C3	8,94427191
100	990-Asal-usul Wilayah Lain	4	1			C3	9,219544457

b. Iterasi Ke-2

Setelah mendapatkan hasil dari perhitungan iterasi awal, selanjutnya dilakukan perhitungan iterasi kedua yang dimana pada proses perhitungan ini akan mendapatkan nilai centroid baru berdasarkan nilai rata-rata pada proses perhitungan dari cluster iterasi-1. Sama seperti halnya iterasi awal pada iterasi kedua seluruh data dimasukkan kedalam cluster terdekat seperti perhitungan dibawah ini:

1. Menentukan Centroid Baru

Tabel 6. Centroid Baru Iterasi -2

Centroid Baru		
Centroid 1	58,8	50,4
Centroid 2	41	17,75
Centroid 3	6,5443038	7,62025316

2. Melakukan Perhitungan dengan jarak pusat cluster baru

Jarak antara buku dengan pusat cluster 1

$$D_{11} = \sqrt{(58,8 - 25)^2 + (50,4 - 32)^2}$$

$$D_{11} = 38,48376281$$

$$D_{12} = \sqrt{(58,8 - 10)^2 + (50,4 - 28)^2}$$

$$D_{12} = 53,6954374$$

$$D_{13} = \sqrt{(58,8 - 83)^2 + (50,4 - 57)^2}$$

$$D_{13} = 25,08385935$$

$$D_{14} = \sqrt{(58,8 - 31)^2 + (50,4 - 17)^2}$$

$$D_{14} = 43,45572459$$

$$D_{15} = \sqrt{(58,8 - 66)^2 + (50,4 - 15)^2}$$

$$D_{15} = 36,12478374$$

Jarak antara buku dengan pusat cluster 2

$$D_{21} = \sqrt{(41 - 25)^2 + (17,75 - 32)^2}$$

$$D_{21} = 21,42574386$$

$$D_{22} = \sqrt{(41 - 10)^2 + (17,75 - 28)^2}$$

$$D_{22} = 32,65061255$$

$$D_{23} = \sqrt{(41 - 83)^2 + (17,75 - 67)^2}$$

$$D_{23} = 57,48532421$$

$$D_{24} = \sqrt{(41 - 31)^2 + (17,75 - 17)^2}$$

$$D_{24} = 10,02808556$$

$$D_{25} = \sqrt{(41 - 66)^2 + (17,75 - 15)^2}$$

$$D_{25} = 25,15079522$$

Jarak antara buku dengan pusat cluster 3

$$D_{31} = \sqrt{(6,5443038 - 25)^2 + (7,62025316 - 32)^2}$$

$$D_{31} = 30,5775208$$

$$D_{32} = \sqrt{(6,5443038 - 10)^2 + (7,6202531628,6 - 28)^2}$$

$$D_{32} = 20,67065353$$

$$D_{33} = \sqrt{(6,5443038 - 83)^2 + (7,62025316 - 67)^2}$$

$$D_{33} = 91,01556394$$

$$D34 = \sqrt{(6,5443038 - 31)^2 + (7,62025316 - 17)^2}$$

$$D34 = 26,19276097$$

$$D35 = \sqrt{(6,5443038 - 66)^2 + (7,62025316 - 15)^2}$$

$$D35 = 59,91193933$$

Hasil perhitungan dari keseluruhan data tersebut dengan cluster baru pada iterasi 2 dapat dilihat pada tabel 7 dibawah ini.

Tabel 7. Hasil Perhitungan Pada Iterasi ke-2

No	Kategori Buku	Peminjam Buku	Stok Buku	Cluster			Jarak Terdekat
		(K1)	(K2)	C1	C2	c3	
1	000-Publikasi umum, Informasi Umum	25	32	38,48376281	21,42574386	30,5775208	21,42574386
2	010-Bibliografi	10	28	53,69543742	32,65061255	20,67065353	20,67065353
3	020-Perpustakaan dan Informasi	83	57	25,08385935	57,48532421	91,01556394	25,08385935
4	030-Ensiklopedia dan Buku	31	17	43,45572459	10,02808556	26,19276097	10,02808556
5	040-Biografi	66	15	36,12478374	25,15079522	59,91193933	25,15079522
6	050-Majalah dan Jurnal	46	7	45,24820438	11,8559057	39,46057117	11,8559057
7	060-Asosiasi, Organisasi dan Museum	12	48	46,86149806	41,90539941	40,74663883	40,74663883
8	070-Media Massa, Jurnalisme dan Publikasi	54	57	8,160882305	41,34685599	68,48651327	8,160882305
9	080-Kutipan	6	10	66,48308055	35,84776841	2,441200857	2,441200857
10	090-Manuskrip dan Buku Langka	41	19	36,09432088	1,25	36,28627342	1,25
11	100-Filsafat dan Psikologi	10	5	66,6528319	33,51958383	4,336768715	4,336768715
12	110-Metafisika	6	5	69,63476143	37,25	2,676190066	2,676190066
13	120-Epistimologi	15	12	58,24946352	26,6282275	9,522656174	9,522656174
14	030-Parapsikologi dan Okultisme	7	7	67,57810296	35,65897503	0,769657727	0,769657727
15	040-Pemikiran Filosofis	4	9	68,68041933	38,02055365	2,894336391	2,894336391
16	050-Psikologi	47	50	11,80677771	32,80339159	58,5893019	11,80677771
17	060-Filosofis Logis	12	4	65,90295896	32,09458677	6,547583831	6,547583831
18	070-Etik	3	1	74,52516354	41,5278521	7,509316971	7,509316971
19	180-Filosofi Kuno	5	2	72,36711961	39,29456069	5,828560702	5,828560702
20	190-Filosofi Barat Modern	1	2	75,38832801	42,98909746	7,894716602	7,894716602
...
99	980-Asal-usul Amerika Selatan	10	2	68,73136111	34,77157604	6,597657302	6,597657302
100	990-Asal-usul Wilayah Lain	4	1	73,77940092	40,61480641	7,092336271	7,092336271

3. Melakukan Hasil dari Pengklasteran Iterasi -2

Setelah menghitung jarak data pada nilai centroid baru, Kemudian masuk ke tahapan pengklasterisasian data. Hasil pengklasterisasian dapat dilihat pada tabel 8 dibawah ini

Tabel 8. Hasil Pengklasteran Nilai Cluster Iterasi 2

No	Kategori Buku	Peminjam Buku	Stok Buku	Cluster			Jarak Terdekat
		(K1)	(K2)	C1	C2	C3	
1	000-Publikasi umum, Informasi Umum	25	32		C2		21,42574386
2	010-Bibliografi	10	28	C1		C3	20,67065353
3	020-Perpustakaan dan Informasi	83	57		C2		25,08385935
4	030-Ensiklopedia dan Buku	31	17		C2		10,02808556
5	040-Biografi	66	15		C2		25,15079522
6	050-Majalah dan Jurnal	46	7			C3	11,8559057
7	060-Asosiasi, Organisasi dan Museum	12	48			C3	40,74663883

No	Kategori Buku	Peminjam Buku	Stok Buku	Cluster			Jarak Terdekat
		(K1)	(K2)	C1	C2	C3	
8	070-Media Massa, Jurnalisme dan Publikasi	54	57	C1			8,160882305
9	080-Kutipan	6	10			C3	2,441200857
10	090-Manuskrip dan Buku Langka	41	19		C2		1,25
11	100-Filsafat dan Psikologi	10	5			C3	4,336768715
12	110-Metafisika	6	5			C3	2,676190066
13	120-Epistimologi	15	12			C3	9,522656174
14	030-Parapsikologi dan Okultisme	7	7			C3	0,769657727
15	040-Pemikiran Filosofis	4	9			C3	2,894336391
16	050-Psikologi	47	50	C1			11,80677771
17	060-Filosofis Logis	12	4			C3	6,547583831
18	070-Etik	3	1			C3	7,509316971
19	180-Filosofi Kuno	5	2			C3	5,828560702
20	190-Filosofi Barat Modern	1	2			C3	7,894716602
...
99	980-Asal-usul Amerika Selatan	10	2			C3	6,597657302
100	990-Asal-usul Wilayah Lain	4	1			C3	7,092336271

Perbandingan proses hasil perhitungan iterasi -3 dengan iterasi -2 menghasilkan nilai yang sama, maka perhitungan ini dihentikan sampai iterasi ke-2.

4. KESIMPULAN

Berdasarkan proses perhitungan Pengelompokkan Buku Untuk Menentukan Minat Baca Pada Metode K-Means menghasilkan jumlah cluster Paling Diminati, Diminati, dan Kurang Diminati yaitu: Pada Kluster 1 (Paling Diminati) memilih 6 kategori buku diantaranya Kategori 020-Perpustakaan dan Informasi, 070-Media Massa, Jurnalisme dan Publikasi, 050-Psikologi, 420-Bahasa Indonesia, 600-Teknologi, 650-Manajemen. Pada Kluster 2 (Diminati) memilih 16 kategori buku diantaranya 000-Publikasi Umum dan Informasi Umum, 030-Ensiklopedia dan Buku, 040-Biografi, 050-Majalah dan Jurnal, 090-Manuskrip dan Buku Langka, 210-Agama Islam, 300-Ilmu Sosial, 320-Ilmu Politik, 330-Ilmu Ekonomi, 410-Bahasa Indonesia, 510-Matematika, 620-Ilmu Teknik, 770-Fotografi dan Foto, 780-Musik, 910-Geografi Umum Perjalanan, 930-Sejarah Dunia Lama. Dan pada kluster 3 (Kurang Diminati) terpilih sebanyak 78 kategori buku. 10 diantaranya 010-Bibliografi, 060-Asosiasi Organisasi dan Museum, 080-Kutipan, 100-Filsafat dan Psikologi, 110-Matematika, 120-Epistimologi, 030-Parapsikologi dan Okultisme, 040-Pemikiran Filosofis, 060-Filosofi Logis, 070-Etik. Metode K-Means dapat diterapkan dalam menyelesaikan permasalahan pada Perpustakaan Daerah Kota Medan dalam hal pengelompokkan buku. Pengaruh data mining dalam menyelesaikan permasalahan ini sangat baik dan didukung dengan proses yang mudah dan hasil yang sangat efektif sehingga pihak-pihak perpustakaan kedepannya akan mengetahui kategori buku apa saja yang paling diminati, diminati, dan Kurang diminati.

REFERENCES

- [1] P. A. Gatto and R. M. Awangga, *Pengelompokan Kedisiplinan Pegawai Berdasarkan Absensi Menggunakan Algoritma K-Means*. Penerbit Buku Pedia, 2023.
- [2] V. Agustin, "Perbandingan Analisis Clustering K-Means Dan K-Medoids Pada Data Penyakit Di Indonesia Tahun 2019," 2021.
- [3] W. M. Putri, E. Asril, and U. L. Kuning, "Analisis Clustering Buku Sebagai Upaya Untuk Meningkatkan Minat Baca Siswa Pada Perpustakaan Sma Negeri 3 Pekanbaru," vol. 2, no. 1, pp. 313–323, 2023.
- [4] D. APRIZA and S. Samsuryadi, "Perbandingan Metode Euclidean Distance Dengan Coefficient Correlation Pada Klasifikasi Penyakit Multiple Sclerosis Lesion ...," 2019, [Online]. Available: https://repository.unsri.ac.id/23207/%0Ahttps://repository.unsri.ac.id/23207/52/RAMA_55201_09121402019_0004027101_01_front_ref.pdf.
- [5] M. A. K. Clustering, "(1) , 2)," vol. 4, no. 2, pp. 300–311, 2022.
- [6] M. Rafi, "Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa," vol. 12, no. 2, pp. 121–129, 2020.
- [7] M. M. Muchlis, I. Fitri, and R. Nuraini, "Rancang Bangun Aplikasi Data Mining pada Penjualan Distro Bloods Berbasis Web menggunakan Algoritma Apriori," *J. JTik (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 5, no. 1, pp. 26–33, 2021.
- [8] M. Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*. Wiley, 2019.
- [9] M. K. Miftahul Huda, *Algoritma Data Mining: Analisis Data Dengan Komputer*. bisakimia.
- [10] C. Penerima, B. Kip, and S. Usanto, "Penerapan Data Mining Dengan Mengimplementasikan Algoritma K- Means Dalam Proses Clustering Untuk Pengelompokan Mahasiswa," vol. 5, no. 1, 2023, doi: 10.47065/bits.v5i1.3411.
- [11] C. Hayashi, K. Yajima, H. H. Bock, N. Ohsumi, Y. Tanaka, and Y. Baba, *Data Science, Classification, and Related Methods: Proceedings of the Fifth Conference of the International Federation of Classification Societies (IFCS-96), Kobe, Japan, March 27--30, 1996*. Springer Japan, 2013.
- [12] S. A. M. M. M. P. M. S. Dr. H. DARMADI, *MEMBACA, YUUK.....! "Strategi Menumbuhkan Minat Baca Pada Anak Sejak Usia Dini."* GUEPEDIA.
- [13] M. Y. Matdoan, "Perbandingan Algoritma K-Means Dan K-Medoids Dalam Pengelompokkan Tingkat Kebahagiaan Provinsi Di Indonesia," *BULLET J. Multidisiplin Ilmu*, vol. 1, no. 05, pp. 830–839, 2022.
- [14] I. Kamila, U. Khairunnisa, and M. Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau," *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 5, no. 1, pp. 119–125, 2019.

- [15] R. D. Astuti, "Analisis Perbandingan Algoritma K-Means Dan K-Medoids Untuk Menerapkan Segmentasi Pelanggan," 2019.
- [16] A. S. Munthe, "KLASIFIKASI BUKU DI PERPUSTAKAAN DENGAN METODE K-MEANS CLUSTERING UNTUK MENGETAHUI BUKU POPULER." Universitas Nasional, 2023.
- [17] S. Putri Aulia Syam, C. Dwi Suhendra, and L. F. Marini, "Penerapan K-Means Clustering dalam Menentukan Bidang Magang Mahasiswa Jurusan Teknik Informatika Universitas Papua," *Infomatek*, vol. 25, no. 1, pp. 49–54, 2023, doi: 10.23969/infomatek.v25i1.7585.
- [18] A. Salam, D. Adiatma, and J. Zeniarja, "Implementasi Algoritma K-Means Dalam Pengklasteran untuk Rekomendasi Penerima Beasiswa PPA di UDINUS," *JOINS (Journal Inf. Syst.)*, vol. 5, no. 1, pp. 62–68, 2020, doi: 10.33633/joins.v5i1.3350.
- [19] D. Darlinda and J. N. Utamajaya, "Sistem Pendukung Keputusan Penerima Beasiswa Program Indonesia Pintar Menggunakan Metode Algoritma K-Means Clustering," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 2, p. 167, 2022, doi: 10.30865/jurikom.v9i2.3971.
- [20] M. Herviany, S. Putri Delima, T. Nurhidayah, and Kasini, "Comparison of K-Means and K-Medoids Algorithms for Grouping Landslide Prone Areas in West Java Province," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 34–40, 2021.