

Implementasi Algoritma K-Means untuk Pengelompokan Peminjaman Buku Pada Perpustakaan Sekolah

Implementation of K-Means Algorithm for Clustering Books Borrowing in School Libraries

Daud Siburian¹, Sundari Retno Andani², Ika Purnama Sari³
STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Article Info

Genesis Artikel:

Diterima, 25 Juli 2022

Direvisi, 25 Juli 2022

Disetujui, 28 Juli 2022

Kata Kunci:

Implementasi
Data Mining
K-Means
Buku
Perpustakaan

ABSTRAK

Perpustakaan sekolah merupakan salah satu sumber penting dalam upaya mendukung proses peningkatan mutu pendidikan di sekolah. Melalui perpustakaan banyak informasi yang dapat digali dan dimanfaatkan untuk kepentingan pendidikan. Perpustakaan diharapkan dapat memainkan fungsinya sebagai wahana pendidikan, penelitian pelestarian, informasi, dan rekreasi untuk meningkatkan kecerdasan bangsa. Penelitian ini bertujuan untuk mengelompokkan (kluster) peminjaman buku-buku perpustakaan pada SMA Swasta Assisi Pematangsiantar. Data penelitian diperoleh dari perpustakaan sekolah tersebut. Algoritma yang digunakan untuk proses klusterisasi adalah K-Means Clustering yang merupakan salah satu algoritma data mining. Data diolah menggunakan Microsoft Excel dan Rapid Miner 5.3 untuk menentukan nilai centroid dalam 2 cluster yaitu cluster tertinggi dan terendah. Berdasarkan perhitungan manual dengan Microsoft Excel dan pengujian dengan Rapid Miner, penelitian ini menghasilkan nilai yang sama, yakni cluster tertinggi menghasilkan 6 jenis buku diantaranya Matematika, Geografi, Kimia, Pkn, Penjaskes dan Komputer. Sedangkan untuk Cluster terendah sebanyak 6 jenis buku yaitu Bahasa Indonesia, Bahasa Inggris, Biologi, Fisika, Agama dan Seni Budaya. Sehingga dapat disimpulkan bahwa metode K-Means pada penelitian ini dapat mengelompokkan peminjaman buku perpustakaan sekolah dengan baik, mengacu pada perhitungan manual dan pengujian yang memiliki hasil yang sama.

ABSTRACT

The school library is an important resource in an effort to support the process of improving the quality of education in schools. Through the library a lot of information can be extracted and used for educational purposes. The library is expected to play its function as a vehicle for education, research, preservation, information, and recreation to improve the nation's intelligence. This study aims to cluster the borrowing of library books at SMA Assisi Pematangsiantar. The research data was obtained from the school library. The algorithm used for the clustering process is K-Means Clustering which is one of the data mining algorithms. The data was processed using Microsoft Excel and Rapid Miner 5.3 to determine the value of the centroid in 2 clusters, namely the highest and lowest clusters. Based on manual calculations with Microsoft Excel and testing with Rapid Miner, this study resulted in the same value, namely the highest cluster produced 6 types of books including Mathematics, Geography, Chemistry, Civics, Physical Education and Computers. As for the lowest cluster, there are 6 types of books, namely Indonesian, English, Biology, Physics, Religion and Cultural Arts. So it can be concluded that the K-Means method in this study can cluster school library book borrowing well, referring to manual calculations and testing which have the same results.

This is an open access article under the [CC BY-SA](#) license.



Penulis Korespondensi:

Daud Siburian,
Program Studi Teknik Informatika,
STIKOM Tunas Bangsa, Pematangsiantar, Indonesia
Email: daudsiburianstb@gmail.com

1. PENDAHULUAN

Perpustakaan merupakan tempat dimana berbagai koleksi buku dan bahan atau karya cetak dikumpulkan. Walaupun dapat diartikan sebagai koleksi pribadi perseorangan, namun perpustakaan lebih umum dikenal sebagai sebuah koleksi besar yang dibiayai dan dioperasikan oleh sebuah kota atau institusi, dan dimanfaatkan oleh masyarakat yang rata-rata tidak mampu membeli sekian banyak buku atas biaya sendiri. Buku perpustakaan merupakan buku-buku bacaan dan tulisan-tulisan karya tulis yang ada di perpustakaan, yang disediakan sekolah untuk dipinjamkan bagi peserta didik sebagai upaya untuk memenuhi kebutuhan buku-buku referensi dan bahan bacaan dalam proses belajar mengajar. Buku perpustakaan merupakan buku-buku bacaan dan tulisan-tulisan karya tulis yang ada di perpustakaan, yang disediakan sekolah untuk dipinjamkan bagi peserta didik sebagai upaya untuk memenuhi kebutuhan buku-buku referensi dan bahan bacaan dalam proses belajar mengajar. Dalam perpustakaan koleksi harus dikelola dengan baik. Khususnya dalam hal penyajian, sangat perlu diperhatikan dengan baik bagaimana menata serta mengelompokkan koleksi perpustakaan. Hal ini bertujuan untuk mempermudah dalam pengelolaan serta juga memberikan kenyamanan juga kemudahan bagi pengguna perpustakaan agar dapat mencari atau menemukan apa yang dicari.

Banyak cabang ilmu komputer yang mampu memecahkan masalah yang kompleks. Hal ini terbukti dari banyaknya penelitian-penelitian yang sudah dilakukan di bidang komputer untuk menyelesaikan masalah komputasi, seperti sistem pendukung keputusan [1]–[6], bidang jaringan syaraf tiruan [7]–[16], hingga di bidang data mining [17]–[23]. Oleh karena itu pada penelitian ini akan dilakukan pengelompokkan buku perpustakaan sekolah dengan mengambil studi kasus pada SMA Swasta Assisi Pematangsiantar. Algoritma pengelompokkan yang akan digunakan adalah algoritma K-Means yang merupakan salah satu algoritma data mining. Algoritma ini digunakan karena algoritma K-means telah terbukti dan banyak digunakan untuk penelitian-penelitian yang berhubungan dengan pengelompokkan data.

Beberapa penelitian terdahulu yang berhubungan dengan topik ini diantaranya: Penelitian untuk mengelompokkan Laju Pertumbuhan PDRB Kota Surabaya menggunakan Metode K-Means. Data dibagi menjadi 3 Kluster: tinggi, sedang dan rendah. Hasil yang diperoleh terdapat 9 kategori/sektor dengan kluster tinggi, 5 kategori/sektor dengan kluster sedang dan 3 kategori/sektor dengan kluster rendah [19]. Selanjutnya penelitian yang dilakukan untuk mengelompokkan jumlah kasus terkonfirmasi Covid-19 dan jumlah kematian akibat virus ini di Asia Tenggara. Data yang digunakan adalah data statistik negara berdasarkan area kasus COVID-19 yang dikonfirmasi dan tercatat pada laboratorium, serta kematian pada bulan April 2020 dari WHO (Kesehatan Dunia Organisasi). Pada penelitian ini Data dibagi menjadi 3 Kluster: tinggi (C1), sedang (C2) dan rendah (C3). Hasil yang diperoleh terdapat empat negara dengan Kluster level tinggi (C1), satu negara dengan Kluster level sedang (C2), dan 6 negara dengan Kluster level rendah (C3) [18]. Berikutnya adalah penelitian menggunakan metode K-Means untuk mengelompokkan garis kemiskinan menurut kabupaten/kota di Sumatera Utara. Data yang digunakan adalah data garis kemiskinan menurut kabupaten/kota (rupiah/kapita/bulan) di Provinsi Sumatera Utara tahun 2017-2019. Data bersumber dari Badan Pusat Statistik Sumatera Utara. Pengelompokan tersebut dibagi menjadi 3 kluster yaitu garis kemiskinan kategori tinggi, garis kemiskinan kategori sedang, dan garis kemiskinan kategori rendah. Hasil untuk kategori tinggi terdiri dari 5 kabupaten/kota, kategori sedang terdiri dari 18 kabupaten/kota dan kategori sedang terdiri dari 10 kabupaten/kota [20].

Berdasarkan latar belakang permasalahan yang telah diuraikan, maka dilakukan penelitian ini yang diharapkan dapat mengetahui pengelompokan Buku pada Perpustakaan Sekolah di SMA Swasta Assisi Pematangsiantar.

2. METODE PENELITIAN

2.1. Pengumpulan Data

Dataset penelitian yang digunakan pada penelitian ini adalah data buku perpustakaan yang diperoleh dari sekolah SMA Swasta Assisi Pematangsiantar, yang diperoleh dengan cara :

1. Wawancara

Tujuan dari wawancara ini untuk mendapatkan informasi yang dibutuhkan dalam penelitian ini. Wawancara dilakukan dengan menyampaikan sejumlah pertanyaan kepada narasumber. Adapun narasumber yang diwawancarai yaitu kepala perpustakaan sekolah SMA Swasta Assisi Pematangsiantar.

2. Observasi

Hal ini dilakukan dengan mengamati dan mencatat hal-hal yang berhubungan dengan penelitian serta mengumpulkan dataset penelitian berupa data buku perpustakaan di SMA Swasta Assisi Pematangsiantar tahun 2018-2020.

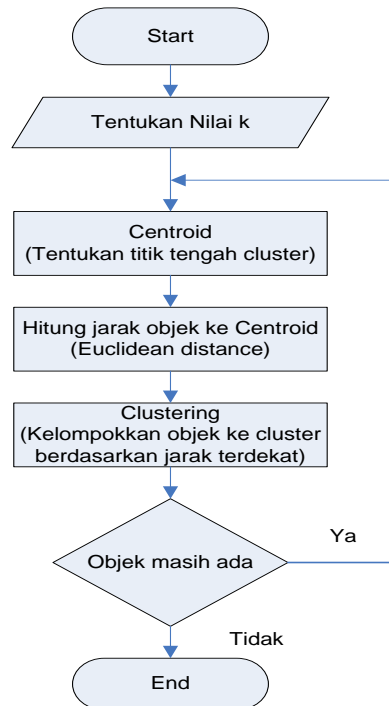
Tabel 1. Dataset Peminjaman Buku Perpustakaan di SMA Swasta Assisi Pematangsiantar Tahun 2018-2020

No	Nama Buku	Tahun		
		2018	2019	2020
1	Bahasa Indonesia	150	140	110
2	Bahasa Inggris	100	120	123
3	Biologi	150	150	132
4	Matematika	140	160	145
5	Geografi	170	148	167
6	Kimia	130	146	154
7	Fisika	110	124	125
8	Pkn	135	136	167

No	Nama Buku	Tahun		
		2018	2019	2020
9	Penjaskes	145	147	186
10	Agama	125	153	128
11	Seni Budaya	133	125	139
12	Komputer	145	163	163

2.2. Flowchart Penelitian

Flowchart penelitian dengan algoritma K-Means pada penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Flowchart Penelitian [24]

Tahapan demi tahapan metode *K-Means* sebagai berikut [25]–[29]:

1. Atur jumlah *cluster* (*k*) pada data set.
2. Atur nilai pusat (*Centroid*).

Pada tahap awal Pengaturan nilai *Centroid* dilakukan secara acak. Pada tahap iterasi digunakan rumus persamaan (1) seperti berikut.

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \tag{1}$$

Keterangan :

V_{ij} = *Centroid* rata-rata *cluster* ke-*i* variabel ke-*j*

N_i = Jumlah anggota *cluster* ke-*i*

i, k = Indeks *cluster*

j = Indeks variabel

X_{kj} = nilai ke-*k* variabel ke-*j* untuk *cluster*

3. Berdasarkan masing-masing *record*, jarak terdekat dihitung dengan *Centroid*.

Beberapa cara yang biasa digunakan untuk melakukan pengukuran jarak data ke pusat *cluster*, diantaranya *Euclidean*, *Manhattan/City Block*, dan *Minkowsky*. Setiap cara memiliki kelebihan dan kekurangan masing-masing. Untuk penulisan pada bab ini, jarak *Centroid* yang digunakan adalah *Euclidean Distance*, dengan rumus seperti dibawah ini:

$$De = \sqrt{(xi - si)^2 + (yi - ti)^2} \tag{2}$$

Keterangan :

De = *Distance Euclidean*

i = Jumlah objek ²

(*x, y*) = Koordinat objek

(*s, t*) = Koordinat *Centroid*

4. Kelompokkan objek berdasarkan jarak ke *Centroid* terdekat
5. Ulangi langkah ke-3 hingga langkah ke-4, lakukan iterasi hingga *Centroid* bernilai optimal

3. HASIL DAN PEMBAHASAN

3.1. Perhitungan manual Algoritma K-Means Clustering

1. Menentukan Data Kluster

Data yang akan di *cluster* adalah data Jumlah Peminjam Buku Pada Perpustakaan SMA Swasta Assisi Pematangsiantar (tabel 1)

2. Menentukan Nilai k Jumlah *Cluster*

Jumlah *cluster* sebanyak 2 *cluster*. *Cluster* yang dibentuk yaitu *cluster* tinggi (C1) dan *cluster* rendah (C2).

3. Menentukan Nilai Centroid (Pusat *Cluster*)

Penentuan pusat *cluster* awal ditentukan secara random yang diambil dari data yang ada dalam range. Adapun nilai untuk *cluster* tinggi (*cluster* 1) diambil dari nilai tertinggi yang terdapat pada tabel 1 dan nilai untuk *cluster* terendah (*cluster* 2) diambil dari nilai terendah yang terdapat pada tabel 1. Tabel centroid data dapat dilihat pada tabel 2 berikut.

Tabel 2. Centroid Data Awal

	2018	2019	2020
Cluster 1	170	163	186
Cluster 2	100	120	110

4. Menghitung Jarak Setiap Data Terhadap Centroid (Pusat *Cluster*)

Setelah data nilai pusat *cluster* awal ditentukan, maka langkah selanjutnya adalah menghitung jarak masing-masing data terhadap pusat *cluster*. Proses pencarian jarak terpendek pada iterasi 1 dapat dilihat pada perhitungan dan tabel berikut:

$$D_{B1,c1} = \sqrt{\frac{(150 - 170)^2 + (140 - 163)^2 + (110 - 186)^2}{}} = 81.88406$$

$$D_{B1,c2} = \sqrt{\frac{(150 - 100)^2 + (140 - 120)^2 + (110 - 110)^2}{}} = 53.85165$$

$$D_{B2,c1} = \sqrt{\frac{(100 - 170)^2 + (120 - 163)^2 + (123 - 186)^2}{}} = 103.5278$$

$$D_{B2,c2} = \sqrt{\frac{(100 - 100)^2 + (120 - 120)^2 + (123 - 110)^2}{}} = 13$$

$$D_{B3,c1} = \sqrt{\frac{(150 - 170)^2 + (150 - 163)^2 + (132 - 186)^2}{}} = 59.03389$$

$$D_{B3,c2} = \sqrt{\frac{(150 - 100)^2 + (150 - 120)^2 + (132 - 110)^2}{}} = 62.32175$$

$$D_{B4,c1} = \sqrt{\frac{(140 - 170)^2 + (160 - 163)^2 + (140 - 186)^2}{}} = 50.89204$$

$$D_{B4,c2} = \sqrt{\frac{(140 - 100)^2 + (160 - 120)^2 + (140 - 110)^2}{}} = 66.52067$$

$$D_{B5,c1} = \sqrt{\frac{(170 - 170)^2 + (148 - 163)^2 + (167 - 186)^2}{}} = 24.20744$$

$$D_{B5,c2} = \sqrt{\frac{(170 - 100)^2 + (148 - 120)^2 + (167 - 110)^2}{}} = 94.51455$$

$$D_{B6,c1} = \sqrt{\frac{(130 - 170)^2 + (146 - 163)^2 + (154 - 186)^2}{}} = 53.97222$$

$$D_{B6,c2} = \sqrt{\frac{(130 - 100)^2 + (146 - 120)^2 + (154 - 110)^2}{}} = 59.26213$$

$$D_{B7,c1} = \sqrt{\frac{(110 - 170)^2 + (124 - 163)^2}{(125 - 186)^2}} = 94.03191$$

$$D_{B7,c2} = \sqrt{\frac{(110 - 100)^2 + (124 - 120)^2}{(125 - 110)^2}} = 18.46619$$

$$D_{B8,c1} = \sqrt{\frac{(135 - 170)^2 + (136 - 163)^2}{(167 - 186)^2}} = 48.11445$$

$$D_{B8,c2} = \sqrt{\frac{(135 - 100)^2 + (136 - 120)^2}{(167 - 110)^2}} = 68.775$$

$$D_{B9,c1} = \sqrt{\frac{(145 - 170)^2 + (147 - 163)^2}{(186 - 186)^2}} = 29.68164$$

$$D_{B9,c2} = \sqrt{\frac{(145 - 100)^2 + (147 - 120)^2}{(186 - 110)^2}} = 92.358$$

$$D_{B10,c1} = \sqrt{\frac{(125 - 170)^2 + (153 - 163)^2}{(126 - 186)^2}} = 74.08779$$

$$D_{B10,c2} = \sqrt{\frac{(125 - 100)^2 + (153 - 120)^2}{(126 - 110)^2}} = 45.14421$$

$$D_{B11,c1} = \sqrt{\frac{(133 - 170)^2 + (125 - 163)^2}{(139 - 186)^2}} = 70.86607$$

$$D_{B11,c2} = \sqrt{\frac{(133 - 100)^2 + (125 - 120)^2}{(139 - 110)^2}} = 44.21538$$

$$D_{B12,c1} = \sqrt{\frac{(145 - 170)^2 + (163 - 163)^2}{(163 - 186)^2}} = 33.97058$$

$$D_{B12,c2} = \sqrt{\frac{(145 - 100)^2 + (163 - 120)^2}{(163 - 110)^2}} = 81.74962$$

Hasil dari keseluruhan perhitungan dapat dilihat pada tabel 3 berikut:

Tabel 3. Hasil Perhitungan Jarak Pusat Cluster Iterasi 1

No	Nama Buku	Tahun			C1	C2	Jarak Terpendek
		2018	2019	2020			
1	Bahasa Indonesia	150	140	110	81.88406	53.85165	53.85164807
2	Bahasa Inggris	100	120	123	103.5278	13	13
3	Biologi	150	150	132	59.03389	62.32175	59.03388857
4	Matematika	140	160	145	50.89204	66.52067	50.8920426
5	Geografi	170	148	167	24.20744	94.51455	24.20743687
6	Kimia	130	146	154	53.97222	59.26213	53.97221507
7	Fisika	110	124	125	94.03191	18.46619	18.46618531
8	Pkn	135	136	167	48.11445	68.775	48.11444689
9	Penjaskes	145	147	186	29.68164	92.358	29.68164416
10	Agama	125	153	128	74.08779	45.14421	45.14421336
11	Seni Budaya	133	125	139	70.86607	44.21538	44.21538194
12	Komputer	145	163	163	33.97058	81.74962	33.9705755

5. Menentukan Posisi Cluster

Penentuan masing-masing data hasil cluster Jumlah Peminjam Buku pada Perpustakaan SMA Assisi Swasta Pematangsiantar berdasarkan jarak minimum data terhadap pusat cluster.

Tabel 4. Hasil Perhitungan Jarak Pusat Cluster Iterasi 1

No	Nama Buku	Jarak Terpendek	C1	C2
1	Bahasa Indonesia	53.85164807		1
2	Bahasa Inggris	13		1

No	Nama Buku	Jarak Terpendek	C1	C2
3	Biologi	59.03388857	1	
4	Matematika	50.8920426	1	
5	Geografi	24.20743687	1	
6	Kimia	53.97221507	1	
7	Fisika	18.46618531		1
8	Pkn	48.11444689	1	
9	Penjaskes	29.68164416	1	
10	Agama	45.14421336		1
11	Seni Budaya	44.21538194		1
12	Komputer	33.9705755	1	

Proses K-Means akan terus ber iterasi sampai pengelompokan data sama dengan pengelompokan data iterasi sebelumnya. Dengan kata lain, proses akan terus melakukan iterasi sampai data pada iterasi terakhir sama dengan iterasi sebelumnya.

6. Menghitung *centroid* baru menggunakan hasil pada masing-masing *cluster*

Setelah didapatkan hasil jarak dari setiap objek pada iterasi ke-1 maka lanjut ke iterasi ke-2 pada perhitungan dan tabel dibawah ini :

$$D_{2018,c1} = \frac{150+140+170+130+135+145+145}{7} = 145$$

$$D_{2019,c1} = \frac{150+160+148+146+136+147+163}{7} = 150$$

$$D_{2020,c1} = \frac{132+145+167+154+167+186+163}{7} = 159,14$$

Tabel 5. *Centroid* Baru Iterasi 1

	2018	2019	2020
<i>Cluster 1</i>	145	150	159.1429
<i>Cluster 2</i>	123.6	132.4	125

7. Melakukan kembali langkah ke 4 sampai 6

Jika nilai *centroid* hasil iterasi dengan nilai *centroid* sebelumnya bernilai sama serta posisi *cluster* tidak mengalami perubahan maka proses iterasi berhenti. Namun jika nilai *centroid* tidak sama serta posisi data masih berubah maka proses iterasi berlanjut pada iterasi berikutnya. Berikut tabel hasil *centroid* baru :

Tabel 6. *Centroid* Baru

	2018	2019	2020
<i>Cluster 1</i>	145	150	159.1429
<i>Cluster 2</i>	123.6	132.4	125

Hasil dari keseluruhan perhitungan dapat dilihat pada tabel 7 sebagai berikut :

Tabel 7. Hasil Perhitungan Jarak Pusat *Cluster* Iterasi 2

No	Nama Buku	Tahun			C1	C2
		2018	2019	2020		
1	Bahasa Indonesia	150	140	110	50.39862	31.30048
2	Bahasa Inggris	100	120	123	65.04849	26.73425
3	Biologi	150	150	132	27.59954	32.49185
4	Matematika	140	160	145	18.02832	37.82486
5	Geografi	170	148	167	26.28183	64.50054
6	Kimia	130	146	154	16.35387	32.66374
7	Fisika	110	124	125	55.37811	15.98499
8	Pkn	135	136	167	18.91388	43.6683
9	Penjaskes	145	147	186	27.02418	66.27307
10	Agama	125	153	128	37.13324	20.86432
11	Seni Budaya	133	125	139	34.2744	18.41521
12	Komputer	145	163	163	13.56015	53.27589

Tabel 8. Hasil *Cluster* iterasi 2

No	Nama Buku	Tahun			Jarak Terpendek	C1	C2
		2018	2019	2020			
1	Bahasa Indonesia	150	140	110	31.30048		1
2	Bahasa Inggris	100	120	123	26.73425		1
3	Biologi	150	150	132	32.49185	1	
4	Matematika	140	160	145	37.82486	1	
5	Geografi	170	148	167	64.50054	1	
6	Kimia	130	146	154	32.66374	1	
7	Fisika	110	124	125	15.98499		1
8	Pkn	135	136	167	43.6683	1	
9	Penjaskes	145	147	186	66.27307	1	
10	Agama	125	153	128	20.86432		1
11	Seni Budaya	133	125	139	18.41521		1

No	Nama Buku	Tahun			Jarak Terpendek	C1	C2
		2018	2019	2020			
12	Komputer	145	163	163	53.27589	1	

Tabel 9. Centroid Baru

	2018	2019	2020
Cluster 1	125	127.1429	138.4286
Cluster 2	123.6	132.4	125

Tabel 10. Hasil Perhitungan Jarak Pusat Cluster Iterasi 3

No	Nama Buku	Tahun			C1	C2
		2018	2019	2020		
1	Bahasa Indonesia	150	140	110	39.98112	31.30048
2	Bahasa Inggris	100	120	123	30.23345	26.73425
3	Biologi	150	150	132	34.47862	32.49185
4	Matematika	140	160	145	36.71206	37.82486
5	Geografi	170	148	167	57.23938	64.50054
6	Kimia	130	146	154	24.96119	32.66374
7	Fisika	110	124	125	20.37656	15.98499
8	Pkn	135	136	167	31.54006	43.6683
9	Penjaskes	145	147	186	55.29328	66.27307
10	Agama	125	153	128	27.88094	20.86432
11	Seni Budaya	133	125	139	8.301709	18.41521
12	Komputer	145	163	163	47.84861	53.27589

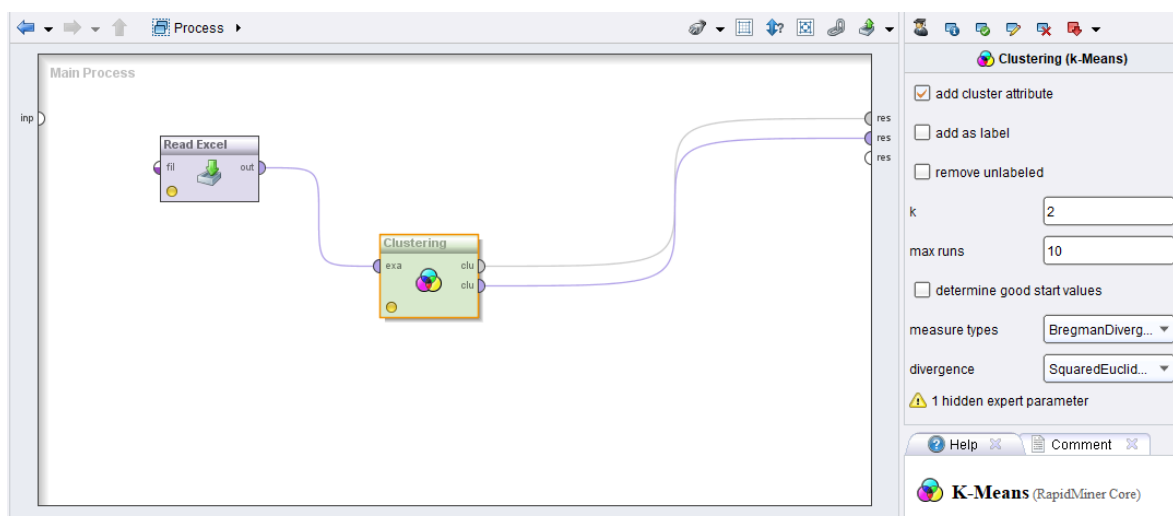
Tabel 11. Hasil Cluster Iterasi 3

No	Nama Buku	Tahun			Jarak Terpendek	C1	C2
		2018	2019	2020			
1	Bahasa Indonesia	150	140	110	31.30047923		1
2	Bahasa Inggris	100	120	123	26.7342477		1
3	Biologi	150	150	132	32.49184513		1
4	Matematika	140	160	145	36.71206219	1	
5	Geografi	170	148	167	57.23938276	1	
6	Kimia	130	146	154	24.96119437	1	
7	Fisika	110	124	125	15.98499296		1
8	Pkn	135	136	167	31.54006199	1	
9	Penjaskes	145	147	186	55.29328114	1	
10	Agama	125	153	128	20.86432362		1
11	Seni Budaya	133	125	139	8.3017087		1
12	Komputer	145	163	163	47.84861331	1	

Perhitungan manual pada data diatas didapatkan hasil akhir yang dimana pada iterasi 2 dan iterasi 3 pengelompokan data yang dilakukan terhadap 2 cluster didapatkan hasil yang sama. Hasil dari iterasi kedua bernilai C1 = 6 dan C2 = 6 pada posisi data tiap cluster. Sehingga posisi cluster pada data tersebut tidak mengalami perubahan lagi maka proses iterasi berhenti sampai iterasi 3.

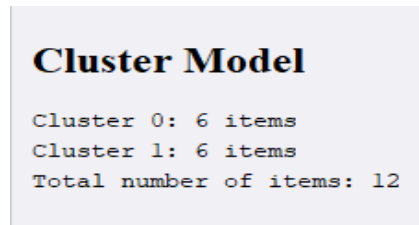
3.2. Pengujian dengan Rapid Miner

Berikut adalah proses pengelompokkan dan hasil dari metode K-Means yang dilakukan dengan Rapid miner. Tahapan pertama dengan meng klik Clustering and Segmentation lalu pilih K-Means dapat dilihat pada gambar 4.9. Lalu hubungkan antara read excel dengan clustering seperti gambar 2 berikut.



Gambar 2. Proses K-Means dengan Rapid Miner

Gambar 2 menjelaskan proses pengelompokan atau pengklusteran metode K-means dengan menggunakan Rapid Miner yang diawali dengan meng import data excel peminjaman buku-buku perpustakaan pada SMA Swasta Assisi Pematangsiantar, kemudian dilanjutkan dengan pemilihan operator metode K-means untuk pengklusteran nya. Nilai k = 2, measure types yang digunakan adalah *Bregman Divergences*. Setelah itu dihubungkan ke operator Apply Model untuk menerapkan model yang sudah dipelajari atau dilatih. Tujuannya adalah untuk mendapatkan prediksi pada unlabeled data (data testing) yang belum memiliki label. Tahap berikutnya adalah menghubungkan ke operator *Clustering* untuk mengevaluasi kinerja model yang memberikan daftar nilai kriteria kinerja secara otomatis sesuai dengan tugas yang diberikan. Hasil nya dapat dilihat pada gambar 3, gambar 4 dan gambar 5.



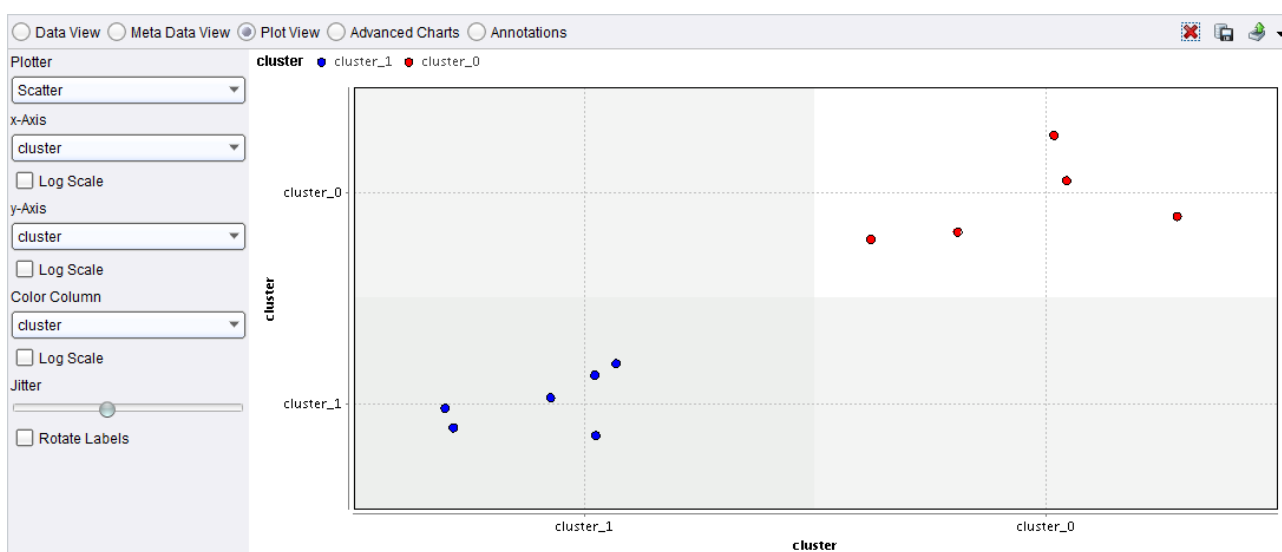
Gambar 3. Hasil Kluster

Berdasarkan gambar 3 dapat dijelaskan bahwa dari dua kluster yang dihasilkan terdapat 6 items untuk *Cluster 0 (C1)*, dan 6 items untuk *Cluster 1 (C2)*. Untuk hasil akhir tabel Centroid dapat dilihat pada gambar 4.

Row No.	NAMA BUKU	id	cluster	2018.0	2019.0	2020.0
1	Bahasa Indc	1	cluster_1	150	140	110
2	Bahasa Ingg	2	cluster_1	100	120	123
3	Biologi	3	cluster_1	150	150	132
4	Matematika	4	cluster_0	140	160	145
5	Geografi	5	cluster_0	170	148	167
6	Kimia	6	cluster_0	130	146	154
7	Fisika	7	cluster_1	110	124	125
8	Pkn	8	cluster_0	135	136	167
9	Penjaskes	9	cluster_0	145	147	186
10	Agama	10	cluster_1	125	153	128
11	Seni Budaya	11	cluster_1	133	125	139
12	Komputer	12	cluster_0	145	163	163

Gambar 4. Data View Hasil Data Centroid

Sedangkan hasil *plot view Cluster* peminjaman buku-buku perpustakaan pada SMA Swasta Assisi Pematangsiantar dapat dilihat pada gambar 5.



Gambar 5. Plot View Peminjaman Buku-Buku Perpustakaan pada SMA Swasta Assisi Pematangsiantar

Gambar 5 menjelaskan bahwa titik warna biru merupakan kelompok *Cluster_1* (tinggi) yang terdiri dari 6 items dan titik warna merah merupakan kelompok *Cluster_0* (rendah) yang juga terdiri dari 6 items.

4. KESIMPULAN

Hasil yang diperoleh dari metode K-Means *clustering* ke dalam Rapid Miner memiliki nilai yang sama yaitu menghasilkan 2 *cluster*. *Cluster* tertinggi menghasilkan 6 jenis buku diantaranya Matematika, Geografi, Kimia, Pkn, Penjaskes, Komputer. Sedangkan untuk *Cluster* terendah sebanyak 6 yaitu Bahasa Indonesia, Bahasa Inggris, Biologi, Fisika, Agama, Seni Budaya. Hasil yang didapat dari penelitian diharapkan dapat menjadi masukan Pihak Sekolah agar memberi perhatian lebih pada Jumlah Peminjam Buku Pada Perpustakaan SMA Swasta Assisi Pematangsiantar yang terendah berdasarkan *Cluster* yang telah dilakukan.

REFERENSI

- [1] V. V. Sianipar, A. Wanto, and M. Safii, "Decision Support System for Determination of Village Fund Allocation Using AHP Method," *The IJICS (International Journal of Informatics and Computer Science) ISSN*, vol. 4, no. 1, pp. 20–28, 2020.
- [2] R. Simarmata, R. W. Sembiring, R. Dewi, A. Wanto, and E. Desiana, "Penentuan Masyarakat Penerima Bantuan Perbaikan Rumah di Kecamatan Siantar Barat Menggunakan Metode ELECTRE," *Journal of Computer System and Informatics (JoSYC)*, vol. 1, no. 2, pp. 68–75, 2020.
- [3] R. Watrionthos, W. A. Ritonga, A. Rengganis, A. Wanto, and M. Isa Indrawan, "Implementation of PROMETHEE-GAIA Method for Lecturer Performance Evaluation," *Journal of Physics: Conference Series*, vol. 1933, no. 1, p. 012067, 2021.
- [4] S. R. Ningsih, D. Hartama, A. Wanto, I. Parlina, and Solikhun, "Penerapan Sistem Pendukung Keputusan Pada Pemilihan Objek Wisata di Simalungun," in *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 2019, pp. 731–735.
- [5] N. Nasution, G. W. Bhawika, A. Wanto, N. L. W. S. R. Ginantra, and T. Afriliansyah, "Smart City Recommendations Using the TOPSIS Method," *IOP Conference Series: Materials Science and Engineering*, vol. 846, no. 1, pp. 1–6, 2020.
- [6] R. A. Hutasoit, S. Solikhun, and A. Wanto, "Analisa Pemilihan Barista dengan Menggunakan Metode TOPSIS (Studi Kasus: Mo Coffee)," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 2, no. 1, pp. 256–262, 2018.
- [7] I. M. Muhamad, S. A. Wardana, A. Wanto, and A. P. Windarto, "Algoritma Machine Learning untuk penentuan Model Prediksi Produksi Telur Ayam Petelur di Sumatera," vol. 1, no. 4, pp. 126–134, 2022.
- [8] M. Mahendra, R. C. Telaumbanua, A. Wanto, and A. P. Windarto, "Akurasi Prediksi Ekspor Tanaman Obat, Aromatik dan Rempah-Rempah Menggunakan Machine Learning," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 2, no. 6, pp. 207–215, 2022.
- [9] R. Puspadini, A. Wanto, and N. Arminarahmah, "Penerapan ML dengan Teknik Bayesian Regulation untuk Peramalan," *Journal of Computer System and Informatics (JoSYC)*, vol. 3, no. 3, pp. 147–155, 2022.
- [10] N. L. W. S. R. Ginantra, A. D. GS, S. Andini, and A. Wanto, "Pemanfaatan Algoritma Fletcher-Reeves untuk Penentuan Model Prediksi Harga Nilai Ekspor Menurut Golongan SITC," *Building of Informatics, Technology and Science (BITS)*, vol. 3, no. 4, pp. 679–685, 2022.
- [11] N. Arminarahmah, S. D. Rizki, O. A. Putra, U. Islam, K. Muhammad, and A. Al, "Performance Analysis and Model Determination for Forecasting Aluminum Imports Using the Powell-Beale Algorithm," *IJISTECH (International Journal of Information System & Technology)*, vol. 5, no. 5, pp. 624–632, 2022.
- [12] N. L. W. S. R. Ginantra *et al.*, "Performance One-step secant Training Method for Forecasting Cases," *Journal of Physics: Conference Series*, vol. 1933, no. 1, pp. 1–8, 2021.
- [13] A. Perdana, S. Defit, and A. Wanto, "Optimalisasi Parameter dengan Cross Validation dan Neural Back-propagation Pada Model Prediksi Pertumbuhan Industri Mikro dan Kecil," *Jurnal Sistem Informasi Bisnis*, vol. 01, no. 11, pp. 34–42, 2021.
- [14] N. L. W. S. R. Ginantra, M. A. Hanafiah, A. Wanto, R. Winanjaya, and H. Okprana, "Utilization of the Batch Training Method for Predicting Natural Disasters and Their Impacts," *IOP Conf. Series: Materials Science and Engineering*, vol. 1071, no. 1, p. 012022, 2021.
- [15] A. Wanto, S. Defit, and A. P. Windarto, "Algoritma Fungsi Pelatihan pada Machine Learning berbasis ANN untuk Peramalan Fenomena Bencana," *RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 2, pp. 254–264, 2021.
- [16] V. V. Utari, A. Wanto, I. Gunawan, and Z. M. Nasution, "Prediksi Hasil Produksi Kelapa Sawit PTPN IV Bahjambi Menggunakan Algoritma Backpropagation," *Journal of Computer System and Informatics (JoSYC)*, vol. 2, no. 3, pp. 271–279, 2021.
- [17] N. Arminarahmah, A. D. GS, G. W. Bhawika, M. P. Dewi, and A. Wanto, "Mapping the Spread of Covid-19 in Asia Using Data Mining X-Means Algorithms," *IOP Conf. Series: Materials Science and Engineering*, vol. 1071, no. 012018, pp. 1–7, 2021.
- [18] J. Hutagalung, N. L. W. S. R. Ginantra, G. W. Bhawika, W. G. S. Parwita, A. Wanto, and P. D. Panjaitan, "COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm," *Journal of Physics: Conference Series*,

- vol. 1783, no. 1, p. 012027, 2021.
- [19] N. A. Febriyati, A. D. GS, and A. Wanto, "GRDP Growth Rate *Clustering* in Surabaya City uses the K- Means Algorithm," *International Journal of Information System & Technology*, vol. 3, no. 2, pp. 276–283, 2020.
- [20] M. A. Hanafiah, A. Wanto, and P. B. Indonesia, "Implementation of Data Mining Algorithms for Grouping Poverty Lines by District/City in North Sumatra," *International Journal of Information System & Technology*, vol. 3, no. 2, pp. 315–322, 2020.
- [21] T. H. Sinaga, A. Wanto, I. Gunawan, S. Sumarno, and Z. M. Nasution, "Implementation of Data Mining Using C4.5 Algorithm on Customer Satisfaction in Tirta Lihou PDAM," *Journal of Computer Networks, Architecture, and High-Performance Computing*, vol. 3, no. 1, pp. 9–20, 2021.
- [22] A. Wanto *et al.*, *Data Mining : Algoritma dan Implementasi*. Yayasan Kita Menulis, 2020.
- [23] W. T. C. Gultom, A. Wanto, I. Gunawan, M. R. Lubis, and I. O. Kirana, "Application of The Levenberg Marquardt Method In Predict The Amount of Criminality in Pematangsiantar City," *Journal of Computer Networks, Architecture, and High-Performance Computing*, vol. 3, no. 1, pp. 21–29, 2021.
- [24] Z. S. Younus *et al.*, "Content-based image retrieval using PSO and k-means clustering algorithm," *Arabian Journal of Geosciences*, vol. 8, no. 8, pp. 6211–6224, 2015.
- [25] E. Prasetyo, *Data Mining: Konsep dan Aplikasi menggunakan Matlab*. Yogyakarta: Andi Offset, 2012.
- [26] D. T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition*. New Jersey: John Wiley & Sons, 2005.
- [27] R. Primartha, *Belajar Machine Learning Teori dan Praktik*. Bandung: Informatika Bandung, 2018.
- [28] T. Khotimah, "Pengelompokan Surat dalam Al Qur'an menggunakan Algoritma K-Means," *Jurnal Simetris*, vol. 5, no. 1, pp. 83–88, 2014.
- [29] I. Parlina, A. P. Windarto, A. Wanto, and M. R. Lubis, "Memanfaatkan Algoritma K-Means dalam Menentukan Pegawai yang Layak Mengikuti Assessment Center untuk *Clustering* Program SDP," *CESS (Journal of Computer Engineering System and Science)*, vol. 3, no. 1, pp. 87–93, 2018.