

# Penerapan Data Mining Untuk Rekomendasi Bidang Studi Menggunakan Algoritma *K-Medoids* Pada SMA N 9 Kota Jambi

Armalinda Gustirani<sup>1</sup>, Jasmir<sup>1</sup>, Fachruddin<sup>2</sup>

<sup>1,2,3</sup> Fakultas Ilmu Komputer, Magister Sistem Informasi, Universitas Dinamika Bangsa, Jambi, Indonesia

Email: <sup>1</sup>[armalindag7@gmail.com](mailto:armalindag7@gmail.com), <sup>2</sup>[jasmir@unama.ac.id](mailto:jasmir@unama.ac.id), <sup>3</sup>[fachruddin.stikom@gmail.com](mailto:fachruddin.stikom@gmail.com)

Email Penulis Korespondensi: [armalindag7@gmail.com](mailto:armalindag7@gmail.com)

Artikel Info :

Artikel History :

Submitted : 06-09-2024

Accepted : 25-09-2024

Published : 30-09-2024

**Kata Kunci:**

Pendidikan; *Data Mining*; *K-Medoids*; Studi ; *RapidMiner*

**Keywords:**

*Education*, *Data Mining*, *K-Medoids*, *e Studies*, *RapidMiner*

**Abstrak**– Pendidikan adalah proses yang mengembangkan seseorang dalam hal cara berpikir, perilaku, kepribadian, bahasa, dan perannya dalam masyarakat. Penelitian ini bertujuan menerapkan data mining dalam merekomendasikan bidang studi di perguruan tinggi dengan menggunakan algoritma *K-Medoids*. Dari penelitian yang dilakukan peneliti merekomendasikan 8 *cluster*, yang di hitung secara manual dan menggunakan *RapidMiner*. Hasil yang didapatkan secara manual pada *cluster* 1 terdapat 11 siswa yang masuk Bidang Pertanian dan Peternakan, *cluster* 2 terdapat 65 siswa yang masuk Bidang Olahraga, *cluster* 3 terdapat 29 siswa yang masuk Bidang Kesehatan, *cluster* 4 terdapat 7 siswa yang masuk Bidang Saintek, *cluster* 5 terdapat 30 siswa yang masuk Bidang Pendidikan, *cluster* 6 terdapat 39 siswa yang masuk Bidang Ekonomi, *cluster* 7 terdapat 17 siswa yang masuk Bidang Agama, *cluster* 8 terdapat 27 siswa yang masuk Bidang Seni. Sedangkan menggunakan *RapidMiner* pada *cluster* 1 terdapat 32 siswa, *cluster* 2 terdapat 36 siswa, *cluster* 3 terdapat 26 siswa, *cluster* 4 terdapat 17 siswa, *cluster* 5 terdapat 30 siswa, *cluster* 6 terdapat 31 siswa, *cluster* 7 terdapat 39 siswa, *cluster* 8 terdapat 14 siswa. Hasil dari penerapan algoritma *K-Medoids* diharapkan dapat membantu siswa dalam menentukan bidang studi di perguruan tinggi.

**Abstract**– Education is a process that develops a person in terms of mindset, attitude, character, language, and role in society. This study aims to apply data mining in recommending fields of study in universities using the *K-Medoids* algorithm. From the research conducted, researchers recommend 8 *clusters*, which is calculated manually and using *RapidMiner*. The results obtained manually in *cluster* 1 showed that there were 11 students who entered the Agriculture and Animal Husbandry Sector, *cluster* 2 had 65 students entering the Sports Sector, *cluster* 3 had 29 students entering the Health Sector, *cluster* 4 has 7 students entering the Science and Technology field, *cluster* 5 has 30 students entering the Education Sector, *cluster* 6 has 39 students entering the Economics field, *cluster* 7 there are 17 students entering the Religion Field, *cluster* 8 there are 27 students entering the Arts Field. Meanwhile using *RapidMiner* in *cluster* 1 there were 32 students, *cluster* 2 has 36 students, *cluster* 3 has 26 students, *cluster* 4 has 17 students, *cluster* 5 has 30 students, *cluster* 6 has 31 students, *cluster* 7 has 39 students, *cluster* 8 has 14 students. Hoped the results of applying the *K-Medoidsss* algorithm can help students determine their field of study in college.

## 1. PENDAHULUAN

Pendidikan adalah proses yang mengembangkan seseorang dalam hal cara berpikir, perilaku, kepribadian, bahasa, dan perannya dalam masyarakat. Melalui pendidikan, individu dapat mengubah kepribadian dan karakter mereka menjadi lebih baik, serta meningkatkan kualitas sumber daya manusia [1]. Pendidikan bagi seseorang sangatlah penting karena dengan pendidikan seseorang dapat mengembangkan kecerdasan secara intelektual serta berkontribusi secara positif dalam masyarakat.

Sistem pendidikan di Indonesia memiliki tiga jalur utama yang melibatkan pendidikan formal dan nonformal. Pendidikan juga terbagi kedalam empat tingkatan yaitu usia dini atau dasar, menengah, atas, dan tinggi [2]. Sesuai kebijakan pemerintah yang mewajibkan pendidikan wajib bagi rakyat Indonesia adalah wajib belajar 12 tahun atau setara dengan sekolah menengah atas. Setelah menyelesaikan tingkat tersebut bagi siswa-siswi yang ingin melanjutkan pendidikannya dapat melanjutkan ke tingkat yang lebih tinggi.

Perguruan tinggi adalah lembaga yang menyelenggarakan pendidikan lanjutan setelah jenjang pendidikan menengah atas atau SMA. Ini sejalan dengan UU No.20 tahun 2003 pasal 19 ayat 1 yang berisi “Perguruan tinggi merupakan jenjang pendidikan setelah pendidikan menengah mencakup program Pendidikan diploma, sarjana, magister, spesialis, dan doktor yang diselenggarakan oleh perguruan tinggi”. Perguruan tinggi memiliki berbagai jenis bidang studi atau jurusan yang ditawarkan [3] [4]. Hal tersebut berdampak pula bagi siswa-siswi kelas 12 di SMA N 9 Kota Jambi yang ingin melanjutkan pendidikannya ke perguruan tinggi.

Berdasarkan wawancara dengan Sukariyati selaku ketua panitia pendaftaran mahasiswa baru, mengatakan bahwa metode yang digunakan selama ini dalam proses merekomendasikan siswa-siswi kelas 12 berfokus pada siswa-siswi yang masuk dalam peringkat 10 besar yang akan diikutkan pada SNBP (Seleksi Nasional Berbasis

Prestasi), dimana siswa-siswi yang tidak masuk dalam proses seleksi tersebut hanya diarahkan untuk mengikuti SNBT (Seleksi Nasional Berbasis Tes) saja. Sehingga metode yang digunakan saat ini belum efektif karna hanya berfokus pada siswa yang mengikuti SNBP saja. Ada faktor lain yang memicu pilihan siswa dalam menentukan bidang studi, dimana siswa mengikuti keinginan orang tua, ikut-ikutan teman atau tidak memiliki pilihan lain tanpa mempertimbangkan minat, bakat, dan kondisi akademik siswa tersebut. Pilihan yang diambil secara tergesa-gesa dapat mengakibatkan terlambat menyadari bahwa bidang studi yang di ambil tidak sesuai.

Metode data mining yang bisa diterapkan dalam menentukan bidang studi bagi calon mahasiswa salah satunya adalah *clustering*. *Clustering* adalah suatu metode yang diterapkan dalam pengelompokkan data, di mana data dalam satu kluster memiliki tingkat kesamaan yang maksimal, sementara kesamaan antara data di kluster yang berbeda adalah minimal [5]. Teknik *clustering* memiliki berbagai macam algoritma diantaranya algoritma *K-Means*, *K-Medoids*, *Fuzzy C-Means*, *Agglomerative*, dan lainnya [6]. Algoritma yang diterapkan di penelitian ini berupa algoritma *K-Medoids*.

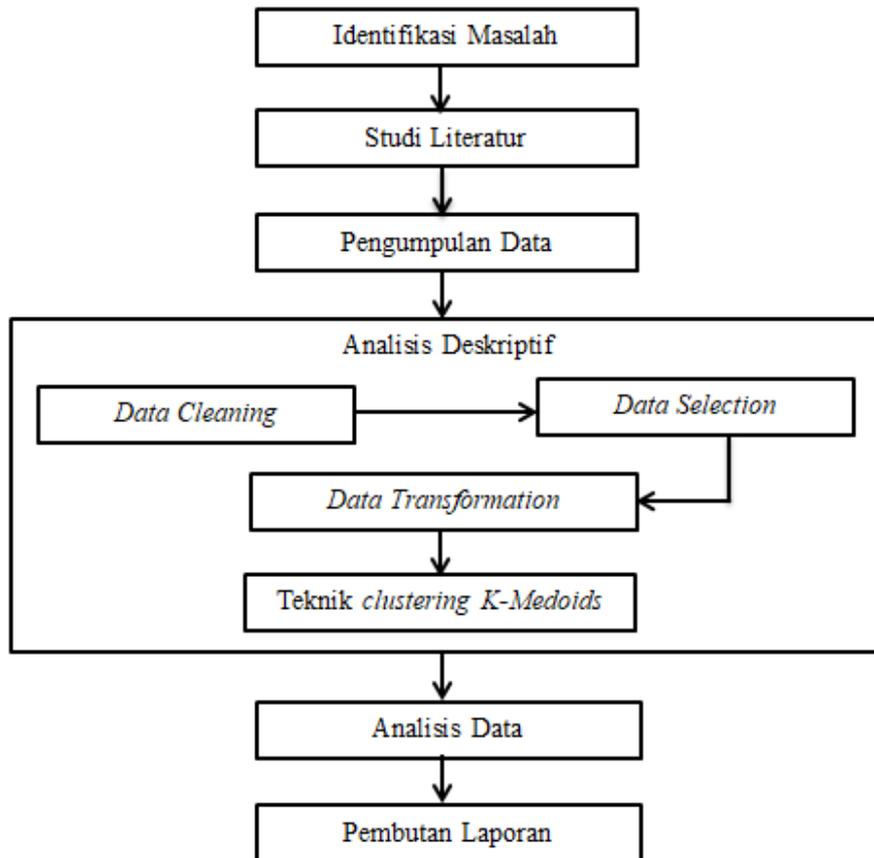
Algoritma *K-Medoids* merupakan algoritma yang digunakan dalam klusterisasi yang berfungsi untuk mengelompokkan atau mengklasifikasikan data ke dalam beberapa *cluster* [7]. Beberapa penelitian terdahulu dengan menggunakan algoritma *K-Medoids* seperti yang dilakukan oleh [8] menunjukkan bahwa algoritma *K-Medoids* bisa digunakan untuk membantu pengelompokkan kelas bagi siswa tunagrahita. Penelitian lainnya yang dilakukan oleh [9] dari proses yang dilakukan didapat hasil bahwa 4 buah *cluster* dengan nilai silhouette coefficient hasil yang terbentuk masuk dalam *cluster* terbaik. Penelitian yang dilakukan oleh [10] diperoleh bahwa algoritma *K-Medoids* bisa menghasilkan 2 kluster yang sesuai, dengan tiga alternatif tergolong di kluster 1 dan dua belas alternatif di kluster 2. Penelitian yang dilakukan oleh [11] menunjukkan nilai *Silhouette Coefficient* terbaik sebesar 0,690754 dengan tiga kluster dan jumlah data sebanyak 15. Penelitian yang dilakukan oleh [12] menunjukkan bahwa kinerja *K-Means* dan *K-Medoids* bervariasi tergantung pada karakteristik dataset. *K-Means* cenderung menghasilkan kelompok yang lebih homogen namun lebih rentan terhadap outlier. Di sisi lain, *K-Medoids* lebih stabil dalam menangani outlier, tetapi mungkin menghasilkan kelompok yang kurang homogen. Pemilihan algoritma *clustering* harus disesuaikan dengan tujuan dan karakteristik spesifik populasi mahasiswa baru program magister.

Berdasarkan uraian diatas, maka peneliti tertarik untuk melakukan penelitian yang berjudul “Penerapan Data Mining Untuk Rekomendasi Pemilihan Bidang Studi Perguruan Tinggi Menggunakan Algoritma *K-Medoids* Pada SMA N 9 Kota Jambi”.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Dalam penyusunan penelitian ini, diperlukan tahapan-tahapan yang jelas. Oleh karena itu, pembahasan mengenai masing-masing tahapan dijabarkan sebagai berikut:



Gambar 1 Kerangka Kerja

Berdasarkan kerangka kerja di atas, berikut penjelasan dari setiap bagian tersebut :

a. Identifikasi masalah

Pada langkah ini, peneliti menentukan dan merumuskan permasalahan yang akan diteliti melalui wawancara, kuesioner, dan pengumpulan data. Tujuannya adalah untuk memahami kesulitan yang dialami siswa dalam memilih bidang studi ke perguruan tinggi. Dengan demikian, peneliti dapat merumuskan rencana kerja dan menentukan data yang diperlukan untuk penelitian ini.

b. Studi Literatur

Pada tahapan ini peneliti mempelajari dan memahami teori-teori melalui *e-book* dan artikel yang relevan untuk menyelesaikan masalah seperti teknik data mining dan *tools* yang akan digunakan.

1. Data Mining

Proses penemuan informasi yang penting dari *database* [13][14].

2. Clustering

Suatu metode yang digunakan dalam mengelompokkan data, di mana data di satu kluster memiliki tingkat kesamaan yang maksimal, sedangkan kemiripan antara data di *cluster* yang berbeda adalah minimum [5].

3. Algoritma *K-Medoids*

Algoritma ini serupa dengan *K-Means*, terutama keduanya merupakan algoritma partisi. Perbedaan utama antara *K-Means* dan *K-Medoids* terletak pada cara penentuan pusat kluster, di mana *K-Medoids* menggunakan objek data sebagai pusat kluster [5][3].

4. Perguruan Tinggi

Berdasarkan UU No.20 tahun 2003 pasal 19 ayat 1 yang berisi “Perguruan tinggi merupakan jenjang pendidikan setelah pendidikan menengah mencakup program Pendidikan diploma, sarjana, magister, spesialis, dan doktor yang diselenggarakan oleh perguruan tinggi”. Bidang studi adalah area pengetahuan yang luas dan mencakup banyak subjek atau topik. Jurusan adalah area spesifikasi dari bidang studi yang lebih tefokus, sedangkan program studi adalah program yang merujuk pada kurikulum spesifikasi yang dirancang untuk memenuhi persyaratan akademik tertentu dalam suatu jurusan.

5. *RapidMiner*

Menurut [15] *RapidMiner* bekerja dengan menganalisis text, mengekstrak informasi dan menggabungkan dengan statistika, AI, dan *database*.

c. Pengumpulan Data

Data yang dipakai data primer dan data sekunder, dengan tujuan untuk mendapatkan data yang diinginkan.

1. Data primer yang digunakan berupa kuesioner yang di sebarakan kepada seluruh anak kelas 12, Serta wawancara dengan pihak SMA Negeri 9 Kota Jambi.
2. Data sekunder yang peneliti peroleh berupa data akademik siswa kelas 12.

d. Analisis Deskriptif

Proses analisis dilakukan dengan beberapa tahap berikut :

1. Data *Cleaning*

Peneliti melakukan persiapan data sebanyak 240 data siswa, setelah itu peneliti melakukan pembersihan data dengan tujuan untuk mengurangi redundansi data.

2. Data *Selection*

Penyeleksian data dilakukan dengan memilih data yang akan dipakai untuk proses data mining.

3. Data *Transformation*

Peneliti mengubah data menjadi format yang sesuai.

4. Teknik *Clustering K-Medoids*

Pada tahap ini dilakukan penerapan metode clustering *K-Medoids*, kemudian dilakukan perhitungan secara manual menggunakan excel.

5. Analisis Data

Pada tahap ini didapat hasil dari perhitungan dengan algoritma *K-Medoids*, yang akan digunakan sebagai rekomendasi bidang studi perguruan tinggi bagi siswa-siswi SMA N 9 Kota Jambi. Hasil analisis perhitungan tersebut akan dilakukan perbandingan dengan hasil perhitungan menggunakan *tools RapidMiner*.

e. Pembuatan Laporan

Pada tahap ini, pembuatan laporan akhir penelitian dilakukan sesuai dengan kerangka yang telah disusun. Laporan tersebut mencakup pendahuluan, landasan teori, metodologi penelitian, analisis, penutup, dan lampiran bukti hasil penelitian.

## 2.2 Metode Analisa Data

a. Proses Pengolahan Data Mining

Tahapan proses data mining sebagai berikut [16] :

1. Data *cleaning*

Langkah untuk menghilangkan data yang sama, mengecek konsistensi, dan memperbaiki kesalahan penulisan.

2. Data *integration*

Menyatukan sejumlah data menjadi data baru yang diperlukan.

3. Data *selection*

Memilih data yang sesuai untuk analisis.

4. Data *transformation*

Menyesuaikan data menjadi format tertentu yang akan digunakan.

5. Data mining

Cara untuk menemukan informasi yang menarik dengan memanfaatkan metode atau algoritma tertentu.

6. *Pattern evaluation*

Hasil yang berupa informasi diperiksa untuk menentukan apakah hipotesis yang ada telah terpenuhi atau tidak.

7. *Knowledge presentation*

Mepresentasikan hasil yang didapat dalam format yang mudah dipahami.

b. Proses Perhitungan Menggunakan *K-Medoids*

Menurut [17] algoritma dari *K-Medoids* sebagai berikut :

1. Inisialisasi data

2. Menentukan jumlah k (jumlah *cluster*) yang akan dibentuk.

3. Memilih *medoids* awal secara sebanyak k dari data

4. Hitung jarak setiap objek (data) dengan kandidat *Medoids* menggunakan persamaan *Euclidian Distance*.

5. Memilih *medoids* baru dan hitung jarak setiap objek.

6. Menghitung total simpangan (S) dengan cara nilai total distance baru dikurang total distance lama. Perhitungan berhenti jika  $S > 0$

7. Lakukan kembali langkah 3-5 sampai tidak ada perubahan pada *Medoids*, kemudian diperoleh klaster dan anggotanya.

Untuk mengukur jarak dapat dilakukan dengan menggunakan *Euclidean Distance* yang dinyatakan dalam Persamaan :

$$d(x,y) = |x-y| \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots (1)$$

Dimana :

d = jarak antara x dan y

x = data pusat klaster.

y = data pada atribut.

i = setiap data

n = jumlah data.

xi = data pada pusat klaster ke i

yi = data pada setiap data ke i

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Analisis Data Mining Menggunakan K-Medoids Clustering

##### 3.1.1 Representasi Data

Digunakan 225 data siswa kelas XII yang sudah di rata-ratakan dan yang mengisi kuesioner. Berikut tabel data siswa.

Tabel 1 Data Siswa SMA N 9 Kota Jambi Kelas 12 T.A 2023/2024

NO	NAMA	Jurusan	Hobi	Minat	Sifat	Bakat	Agama	Bahasa Inggris	Biologi IPA	Sejarah IPS
1	Afiqah Zafirah	IPA	Membaca dan Menulis	Bidang Kesehatan	Optimis	Lainnya	89	81	85	0
2	Aprilia S.Taren	IPA	Traveling	Bidang SAINTEK	Humoris	Lainnya	89	85	84	0
34	Silvina Atsilah	IPA	Traveling	Bidang Agama	Optimis	Aka demik	90	80	81	0
35	Adelia Nurathiyah	IPA	Membaca dan Menulis	Lainnya	Optimis	Musik	90	81	88	0
199	Zahra Nabila	IPS	Menghitung	Bidang Kesehatan	Humoris	Musik	93	80	0	89
200	Adrian Fernando	IPS	Membaca dan Menulis	Bidang Ekonomi	Optimis	Renang	79	72	0	77
224	Tri Cahya Ningtias	IPS	Menyanyi	Bidang Kesehatan	Mandiri	Bemyanyi	82	76	0	79
225	Yunisya Adryani	IPS	Traveling	Bidang Olahraga	Optimis	Lainnya	89	81	0	85

##### 3.1.2 Transformasi Data

Peneliti menggunakan 225 data siswa SMA N 9 Kota Jambi, dengan atribut 20 yang akan dipakai berupa Hobi, Minat, Bakat, Sifat, dan nilai rata-rata mata pelajaran siswa (pendidikan agama, PPKN, bahasa Indonesia, bahasa inggris, matematika umum, sejarah umum, olahraga, seni budaya, fisika IPA, kimia IPA, biologi IPA, matematika minat, ekonomi IPS, sosiologi IPS, sejarah IPS, dan geografi IPS). Setelah itu atribut hobi, minat, bakat, sifat akan di transformasi dengan melakukan inisialisasi terlebih dahulu sehingga data tersebut dapat digunakan, kemudian data nilai rapor di rata-ratakan. Dapat dilihat dibawah ini.

Tabel 2 Inisialisasi

Inisialisasi Hobi			Inisialisasi Minat			Inisialisasi Bakat			Inisialisasi Sifat		
No.	Keterangan Hobi	Inisial	No.	Keterangan Minat	Inisial	No.	Keterangan Bakat	Inisial	No.	Keterangan Sifat	Inisial
1	Membaca dan Menulis	1	2	Bidang Pertanian dan Perternakan	1	3	Music	1	4	Optimis	1
	Menghitung	2		Bidang Olahraga	2		Bernyanyi	2		Humoris	2
	Traveling	3		Bidang Kesehatan	3		Sepak Bola	3		Kreatif	3
	Memasak	4		Bidang SAINTEK	4		Basket	4		Ambisius Dan Kritis	4
	Olahraga	5		Bidang Bahasa	5		Kuliner	5		Egois	5
	Menyanyi	6		Bidang Ekonomi	6		Otomotif	6		Rendah Hati	6
	Menggambar	7		Bidang Agama	7		Bahasa Asing	7		Mandiri	7
	Photography	8		Bidang Pendidikan	8		Renang	8		Mudah Bergaul	8
	Fashion	9		Bidang Seni	9		Akademik	9		Pesimis	9
	lainnya	10		Lainnya	10		Lainnya	10		Lainnya	10

Keterangan tabel :

H	: Hobi	SB	: Seni Budaya
M	: Minat	PJOK	: PJOK
B	: Bakat	MTKP	: Matematika Peminatan
S	: Sifat	FIPA	: Fisika IPA
A	: Agama	KIPA	: Kimia IPA
PPKN	: PPKN	BIPA	: Biologi IPA
BI	: Bahasa Indonesia	EIPS	: Ekonomi IPS
BING	: Bahasa Inggris	SIPS	: Sosiologi IPS
MTKU	: Matematika Umum	GIPS	: Geografi IPS
SI	: Sejarah Indonesia	SJIPS	: Sejarah IPS

Tabel 3 Data Hasil Transformasi

NO	NAMA	H	M	S	B	A	PPKN	MTK U	SI	BI	BING	SB	MTK M	BIPA	FIPA	KIPA	PJ OK	GI PS	SJI PS	SIP S	EI PS
1	Afiqah Zafirah	1	3	1	10	89	89	82	82	88	81	85	83	85	81	81	86	0	0	0	0
2	Aprilia S.Taren	3	4	2	10	89	87	85	86	88	85	92	85	84	82	83	87	0	0	0	0
34	Silvina Atsilah	3	7	1	9	90	86	81	87	89	80	93	83	81	83	81	86	0	0	0	0
35	Adelia Nurathiyah	1	10	1	1	90	89	85	86	86	81	91	84	88	83	82	87	0	0	0	0
199	Zahra Nabila	2	3	2	1	93	89	86	89	86	80	88	0	0	0	0	88	88	89	83	89
200	Adnan Fernando	1	6	1	8	79	79	80	76	82	72	84	0	0	0	0	78	75	77	77	76
224	Tni Cahya Ningtias	6	3	7	2	82	81	82	82	87	76	88	0	0	0	0	87	82	79	79	81
225	Yunisya Adryani	3	2	1	10	89	88	84	83	87	81	87	0	0	0	0	88	84	85	84	85

### 3.1.3 Perhitungan Manual K-Medoids Clustering

Tahap ini peneliti menentukan jumlah klaster (C8). Selanjutnya memilih *medoids* awal secara acak sebanyak 8 *medoids*.

Tabel 4 Nilai Medoids Awal

NO	H	M	S	B	A	PP KN	M TK U	SI	BI	BI NG	SB	M TK M	BI PA	FI PA	KI PA	PJ O K	GI PS	SJ IP S	SI PS	EI PS
1	1	3	1	10	89	89	82	82	88	81	85	83	85	81	81	86	0	0	0	0
19	3	3	10	1	91	89	84	87	88	88	92	87	82	85	82	89	0	0	0	0
43	1	8	1	10	84	88	83	85	88	83	91	83	81	82	83	88	0	0	0	0
91	3	3	10	10	83	83	82	83	85	81	89	84	86	81	83	86	0	0	0	0
149	3	10	2	2	89	86	84	83	83	77	87	0	0	0	0	85	83	88	80	85
194	3	2	1	8	85	84	82	79	87	76	88	0	0	0	0	84	83	80	80	85
211	10	10	6	10	74	75	81	73	81	68	84	0	0	0	0	77	76	76	73	73
224	6	3	7	2	82	81	82	82	87	76	88	0	0	0	0	87	82	79	79	81

Hitung jarak masing-masing objek ke medoids awal

Perhitungan di lakukan dengan menggunakan persamaan euclidian distance (D) pada persamaan (1), untuk menghitung jarak ke medoids masing-masing cluster pada siswa No.1 sebagai berikut :

Data siswa No.1 : (1,3,1,10,89,89,82,82,88,81,85,83,85,81,81,86,0,0,0,0)

Medoids 1 : (1,3,1,10,89,89,82,82,88,81,85,83,85,81,81,86,0,0,0,0)

Medoids 2 : (3,3,10,1,91,89,84,87,88,88,92,87,82,85,82,89,0,0,0,0)

Medoids 3 : (1,8,1,10,84,88,83,85,88,83,91,83,81,82,83,88,0,0,0,0)

Medoids 4 : (3,3,10,10,83,83,82,83,85,81,89,84,86,81,83,86,0,0,0,0)

Medoids 5 : (3,10,2,2,89,86,84,83,83,77,87,0,0,0,0,85,83,88,80,85)

Medoids 6 : (3,2,1,8,85,84,82,79,87,76,88,0,0,0,0,84,83,80,80,85)

Medoids 7 : (10,10,6,10,74,75,81,73,81,68,84,0,0,0,0,77,76,76,73,73)

Medoids 8 : (6,3,7,2,82,81,82,82,87,76,88,0,0,0,0,87,82,79,79,81)

Perhitungan Euclidian Distance :

$$DM1 = \sqrt{(1-1)^2 + (3-3)^2 + (1-1)^2 + (10-10)^2 + (89-89)^2 + (89-89)^2 + (82-82)^2 + (82-82)^2 + (88-88)^2 + (81-81)^2 + (85-85)^2 + (83-83)^2 + (85-85)^2 + (81-81)^2 + (81-81)^2 + (86-86)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2} \dots\dots(2)$$

$$DM2 = \sqrt{(3-1)^2 + (3-3)^2 + (10-1)^2 + (1-10)^2 + (91-89)^2 + (89-89)^2 + (84-82)^2 + (87-82)^2 + (88-88)^2 + (88-81)^2 + (92-85)^2 + (87-83)^2 + (82-85)^2 + (85-81)^2 + (82-81)^2 + (89-86)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2} \dots\dots(3)$$

dilanjutkan sampai DM8

Perhitungan jarak ke setiap medoids dilanjutkan untuk data No.2 hingga No.225 terhadap 8 medoids yang telah ditetapkan. Jarak (cost) dan klaster untuk semua sampel dapat dilihat di tabel 5.:

**Tabel 5** Hasil Perhitungan Jarak Antara Masing-Masing Data Dengan Medoids (Iterasi ke-1)

NO	COST 1	COST 2	COST 3	COST 4	COST 5	COST 6	COST 7	COST 8	KEDEKATAN	CLUSTER
1	0,000000	18,934624	11,822013	14,001428	235,812468	232,902898	224,597596	231,182698	0,000000	C1
2	11,297787	13,546955	8,747571	13,493702	237,237350	234,465435	226,473221	232,623731	8,747571	C3
3	12,907362	13,299624	12,353137	12,528368	239,990750	237,280172	229,375413	235,369752	12,353137	C3
4	14,802702	20,271162	15,278743	17,448209	232,733152	229,846383	221,054925	227,770059	14,802702	C1
5	12,494799	20,287927	11,742231	16,385359	232,579019	229,795300	220,847549	227,845298	11,742231	C3
6	13,304135	19,873601	14,706461	17,109062	230,837519	227,985526	219,180930	225,915117	13,304135	C1
7	17,273100	11,926441	13,009227	15,899686	240,019083	237,405055	229,643114	235,432453	11,926441	C2
...	...	...	...	...	...	...	...	...	...	...
224	231,182698	233,258740	230,881961	232,352921	16,443844	11,299558	25,559343	0,000000	0,000000	C8
225	236,280511	238,571666	236,112854	238,056212	14,433295	12,221293	37,608510	18,271289	12,221293	C6

Berdasarkan tabel di atas, total kedekatan diperoleh sebesar 3344,515762, yang dihitung dengan menjumlahkan jarak kedekatan. Setelah mendapatkan total kedekatan, dilakukan perhitungan iterasi kedua untuk menghitung

total simpangan (S). Hasil kedekatan yang tercantum di tabel 5 di atas merupakan total jarak lama. Perhitungan iterasi kedua dilakukan dengan menggunakan kandidat *medoids* baru seperti yang ditunjukkan dalam tabel berikut:

Tabel 6 *Medoids Baru (iterasi ke-2)*

NO	H	M	S	B	A	PP KN	MT KU	SI	BI	BI NG	SB	MT KM	BI PA	FI PA	KI PA	PJ O K	GI PS	SJ IP S	SI PS	EI PS
14	1	4	4	9	92	90	88	85	89	85	90	88	85	86	87	89	0	0	0	0
92	2	10	9	10	80	80	86	80	85	84	88	88	87	85	82	85	0	0	0	0
112	5	2	2	10	79	79	78	75	86	76	85	0	0	0	0	86	82	78	76	80
139	10	7	8	5	85	85	83	78	82	77	85	0	0	0	0	84	80	83	83	83
68	5	2	2	1	87	87	79	81	85	75	86	83	82	81	80	87	0	0	0	0
155	5	9	7	10	91	91	85	89	83	78	92	0	0	0	0	90	86	89	82	88
200	1	6	1	8	79	79	80	76	82	72	84	0	0	0	0	78	75	77	77	76
219	3	1	1	5	85	85	84	80	87	78	88	0	0	0	0	84	80	82	80	84

Dengan mengikuti perhitungan yang sama menggunakan nilai *medoids* baru pada iterasi kedua, diperoleh hasil jarak *medoids* yang dapat dilihat pada tabel berikut :

Tabel 7 Hasil Perhitungan Jarak Antara Masing-Masing Data Dengan Medoids (Iterasi Ke-2)

NO	COST 1	COST 2	COST 3	COST 4	COST 5	COST 6	COST 7	COST 8	KEDEKATAN	CLUSTER
1	14,564340	16,760668	229,303990	233,712644	13,555811	239,554336	225,869166	232,760392	13,555811	C5
2	9,355212	16,052414	231,104565	235,257731	18,138357	240,613549	227,752936	234,259002	9,355212	C1
3	9,436101	14,886235	233,959398	238,030838	19,586730	243,231741	230,639546	237,080577	9,436101	C1
4	20,931316	22,769278	226,179221	230,503449	7,944810	236,676066	222,639439	229,654088	7,944810	C5
5	19,129036	19,813127	226,109266	230,330111	13,349157	236,251392	222,566754	229,685872	13,349157	C5
6	21,408410	22,483772	224,321287	228,673304	7,453858	234,840031	220,656022	227,769533	7,453858	C5
7	8,630180	18,371717	234,136968	238,074022	22,033611	243,109523	230,905608	237,184485	8,630180	C1
...	...	...	...	...	...	...	...	...	...	...
224	237,166524	235,406627	14,082613	12,690154	229,172948	22,002727	19,299741	10,682696	10,682696	C8
225	241,928833	240,855891	19,634663	17,649929	235,051143	14,467895	28,021420	11,728598	11,728598	C8

Berdasarkan tabel 7 didapatkan total kedekatan = 3176,822975, sehingga dapat dilakukan perhitungan total simpangan (S) sebagai berikut :

$$S = \text{Total cost baru} - \text{Total cost lama} \dots\dots (4)$$

$$= 3176,822975 - 3344,515762$$

$$= -167,692786$$

Dari perhitungan diatas didapatkan total simpangan atau  $S < 0$ , sehingga perhitungan dilanjutkan hingga didapatkan nilai total simpangan atau  $S > 0$ .

Tabel 8 Hasil Cluster Perhitungan Manual

NO	Nama	CLUSTER								BIDANG STUDI
		C1	C2	C3	C4	C5	C6	C7	C8	
1	Afiqah Zafirah		V							Bidang olahraga
2	Aprilia.S.Taren		V							Bidang olahraga
3	Aulia Putri Ramadani		V							Bidang olahraga
4	Awali Fadhil Aryanto	V								Bidang pertanian dan perternakan
5	Azkiya Al Yasilfia		V							Bidang olahraga
6	Bagus Fijoyanto	V								Bidang pertanian dan perternakan
...	...	...	...	...	...	...	...	...	...	...
224	Tri Cahya Ningtias					V				Bidang pendidikan
225	Yunisya Adryani						V			Bidang ekonomi
JUMLAH		11	65	29	7	30	39	17	27	

## 3.2 Hasil Perhitungan Manual Dan Perhitungan *RapidMiner*

Dalam perhitungan manual dan perhitungan menggunakan alat *RapidMiner*, data yang diproses mencakup 225 siswa dengan 20 atribut. Baik perhitungan manual maupun menggunakan *RapidMiner*, peneliti merekomendasikan 8 kluster. Pada perhitungan manual, pemilihan pusat kluster dilakukan dengan cara secara acak oleh peneliti, sedangkan pada *RapidMiner* diambil secara acak otomatis oleh sistem. Pada perhitungan manual *cluster 1* ada 11 data, *cluster 2* ada 65 data, *cluster 3* ada 29 data, *cluster 4* ada 7 data, *cluster 5* ada 30 data, *cluster 6* ada 39 data, *cluster 7* ada 17 data, *cluster 8* ada 27 data, sedangkan pada perhitungan menggunakan *RapidMiner cluster 1* ada 32 data, *cluster 2* ada 36 data, *cluster 3* ada 26 data, *cluster 4* ada 17 data, *cluster 5* ada 30 data, *cluster 6* ada 31 data, *cluster 7* ada 39 data, *cluster 8* ada 14 data.

## 4. KESIMPULAN

Penelitian ini menggunakan data kelas 12 dengan 225 data dan 20 atribut yaitu : Hobi, Minat, Bakat, Sifat, nilai rata-rata semester 1 sampai 5 mata pelajaran untuk jurusan IPA dan jurusan IPS. Pada perhitungan yang telah dilakukan peneliti merekomendasikan 8 *cluster*, Hasil yang didapatkan secara manual pada *cluster 1* ada 11 data, *cluster 2* ada 65 data, *cluster 3* ada 29 data, *cluster 4* ada 7 data, *cluster 5* ada 30 data, *cluster 6* ada 39 data, *cluster 7* ada 17 data, *cluster 8* ada 27 data. Sedangkan menggunakan *RapidMiner* pada *cluster 1* ada 32 data, *cluster 2* ada 36 data, *cluster 3* ada 26 data, *cluster 4* ada 17 data, *cluster 5* ada 30 data, *cluster 6* ada 31 data, *cluster 7* ada 39 data, *cluster 8* ada 14 data. Dimana pada *cluster 1* direkomendasikan masuk Bidang Pertanian dan Peternakan, *cluster 2* masuk Bidang Olahraga, *cluster 3* masuk Bidang Kesehatan, *cluster 4* masuk Bidang Saintek, *cluster 5* Bidang Pendidikan, *cluster 6* masuk Bidang Ekonomi, *cluster 7* masuk Bidang Agama, *cluster 8* masuk Bidang Seni. Dari hasil perhitungan tersebut peneliti merekomendasikan hasil *cluster* yang menggunakan *RapidMiner*.

## REFERENCES

- [1] A. O. Safitri, V. D. Yuniarti, and D. Rostika, "Upaya Peningkatan Pendidikan Berkualitas di Indonesia: Analisis Pencapaian Sustainable Development Goals (SDGs)," *J. Basicedu*, vol. 6, no. 4, pp. 7096–7106, 2022, doi: 10.31004/basicedu.v6i4.3296.
- [2] D. Shintia, M. Asbari, F. Khairunisa, and N. Azizah, "Rapor Pendidikan Indonesia: Quo Vadis Kualitas Pendidikan Indonesia?," *J. Inf. Syst. Manag.*, vol. 2, no. 6, pp. 18–21, 2023.
- [3] N. P. Hariani, "Pengelompokan Perguruan Tinggi Negeri di Indonesia Menggunakan Metode Hirarki dan *K-Medoidss* dengan Ukuran Jarak Modifikasi Data Campuran," *Asian J. Innov. Entrep.*, pp. 87–98, 2022.
- [4] PP, "UNDANG-UNDANG REPUBLIK INDONESIA NOMOR 20 TAHUN 2003 TENTANG SISTEM PENDIDIKAN NASIONAL," pp. 1–7, 2003.
- [5] S. R. Silondae and S. Sutardi, "PENERAPAN ALGORITMA *K-MEDOIDSS* DALAM PENENTUAN FAKTOR TERBESAR PEMILIHAN JURUSAN DI FAKULTAS TEKNIK UNIVERSITAS HALU OLEO," *semantik*, vol. 7, no. 1, pp. 43–50, 2021.
- [6] R. T. Aldisa, "Data Mining Penentuan Jurusan Siswa Menggunakan Metode Agglomerative Hierarchical Clustering (AHC)," *J. Media Inform. Budidarma*, vol. 7, no. 2, pp. 873–880, 2023.
- [7] J. Faran and R. T. Aldisa, "Penerapan Data Mining Untuk Penjurusan Kelas dengan Menggunakan Algoritma *K-Medoidss*," *Build. Informatics, Technol. Sci.*, vol. 5, no. 2, pp. 543–552, 2023, doi: 10.47065/bits.v5i2.4313.
- [8] H. Sariangseh, W. Wanayumini, and R. Rosnelly, "Penentuan Kelas Menggunakan Algoritma *K Medoidss* Untuk Clustering Siswa Tunagrahita," *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 83, 2021, doi: 10.30865/mib.v5i1.2547.
- [9] A. D. Cipta, "Klasterisasi Perguruan Tinggi Swasta Berdasarkan Minat Siswa Menggunakan Metode *K-Medoidss*," *J. Informatics Commun. Technol.*, vol. 4, no. 2, pp. 21–29, 2023, doi: 10.52661/j\_ict.v4i2.116.
- [10] J. Faran and R. T. Aldisa, "Penerapan Data Mining Untuk Penjurusan Kelas dengan Menggunakan Algoritma *K-Medoidss*," *Build. Informatics, Technol. Sci.*, vol. 5, no. 2, pp. 543–552, 2023.
- [11] B. Wira, A. E. Budianto, and A. S. Wiguna, "Implementasi Metode *K-Medoidss Clustering* Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru Tahun 2018 Di Universitas Kanjuruhan

- Malang,” *RAINSTEK J. Terap. Sains Teknol.*, vol. 1, no. 3, pp. 53–68, 2019, doi: 10.21067/jtst.v1i3.3046.
- [12] J. Faran and R. T. Aldisa, “Perbandingan Algoritma *K-Means* dan *K-Medoidss* Dalam Pengelompokan Kelas Untuk Mahasiswa Baru Program Magister,” *J. Inf. Syst. Res.*, vol. 5, no. 2, pp. 509–519, 2024, doi: 10.47065/josh.v5i2.4753.
- [13] F. A. Sianturi, P. M. Hasuglan, A. Simangunsong, and B. Nandek, *DATA MINING : Teori dan Aplikasi Weka*, 1st ed. Medan: IOCS Publisher, 2019.
- [14] A. M. Siregar and A. Paspabhuana, *DATA MINING: Pengolahan Data Menjadi Informasi dengan RapidMiner*, 1st ed. Jakarta: CV. Kekata Group, 2017.
- [15] H. Rizqifaluthi and M. A. Yaqin, “Process mining Akademik Sekolah menggunakan *RapidMiner*,” *MATICS J. Ilmu Komput. dan Teknol. Inf. (Journal Comput. Sci. Inf. Technol.)*, vol. 10, no. 2, pp. 47–51, 2018.
- [16] F. Marisa, A. L. Maukar, and T. M. Akhriza, *DATA MINING KONSEP DAN PENERAPANNYA*, 1st ed. Yogyakarta: CV. BUDI UTAMA, 2021.
- [17] A. Wanto *et al.*, *DATA MINING : Algoritma dan Implementasi*, 1st ed. Medan: Yayasan Kita Menulis, 2020.