

Terbit : 01 September 2023

# Penerapan Metode K-Means Untuk Mengidentifikasi Karakter Siswa ( Studi Kasus : SMA Methodist-8 Medan)

<sup>1</sup> Richard P Simanjuntak, <sup>2</sup>Ramson Rikson Maruwahal Sijabat, <sup>3</sup>Sardo P.Sipayung

<sup>1</sup>Universitas Tjut Nyak Dhien, <sup>2</sup>Politeknik Ganesha, <sup>3</sup> AMIK MBP

Medan, Indonesia

<sup>1</sup>richsparlin0@gmail.com, <sup>2</sup>ramsonriksonsibt@gmail.com, <sup>3</sup> pinsarsiphom@gmail.com

## ABSTRAK

Berbagai cara dilakukan oleh pihak sekolah dalam pendidikan karakter berdasarkan kemendiknas, termasuk Sekolah SMA Methodist-8 Medan.

Kurangnya pemahaman karakter peserta didik menjadikan visi misi sekolah belum terlihat tercapai sehingga pendidikan karakter di SMA Methodist-8 Medan belum tepat sasaran .

Oleh sebab itu perlu dilakukan pengelompokan karakter peserta didik di SMA Methodist-8 Medan dengan tujuan pihak sekolah mengetahui karakter yang dimiliki peserta didik di sekolah.

Algoritma K-Means digunakan untuk melakukan pengelompokan karakter peserta didik dengan jumlah cluster sebanyak 3 menggunakan delapan atribut karakter yang diteliti yaitu Jujur, disiplin,percaya diri, peduli,kreatif dan bertanggung jawab kerja keras, mandiri, dengan 230 data peserta didik.

Hasil pengujian dengan 230 data menggunakan Bahasa R menghasilkan jumlah peserta didik yang berkarakter Cukup 124 peserta didik dengan nilai rata-rata centroid (0,795) dengan karakter bertanggungjawab dan kreatif.

Dari hasil penelitian ini diharapkan Karakter Sangat Baik dapat dikembangkan dengan cara sekolah terus melakukan kebiasaan-kebiasaan yang melibatkan peserta didik agar karakter dalam diri peserta didik dapat terlihat sedangkan untuk karakter Tanggungjawab dan kreatif agar tidak lemah atau cukup maka pihak sekolah selalu memberikan bimbingan kepada peserta didik dan memberi contoh kebiasaan-kebiasaan baik serta kegiatan yang dapat diikuti peserta didik disekolah

Kata Kunci: Data mining, K-Means, Pendidikan Karakter

## PENDAHULUAN

Pendidikan tidak hanya dilaksanakan di sekolah karena bisa dilakukan di rumah terutama dalam lingkungan keluarga. Keluarga menyediakan pembentukan karakter anak hingga dewasa. Keluarga juga memiliki peran keluarga dalam pembentukan kepribadian, yang mana hal ini seringkali kurang mudah dibentuk di lingkungan sekolah. Pembelajaran anak terhadap macam-macam norma yang berlaku di rumah dipertanggung jawabkan oleh peran orang tua dalam mendidik anak. Pendidikan karakter dapat dilakukan disemua jenjang pendidikan di sekolah namun pada jenjang sekolah menengah pendidikan karakter harus lebih diutamakan , karena perkembangan peserta didik berada pada jenjang tersebut. Hal tersebut menjadi alasan dasar sekolah sebagai sarana pendidikan formal dijadikan wadah pembentukan karakter anak sehingga di sekolah guru harus lebih intensif dalam membentuk karakter peserta didiknya agar anak-anak tersebut mampu membentuk Indonesia sebagai bangsa yang berkarakter di masa yang akan datang.

Berbagai cara dilakukan oleh pihak sekolah dalam pendidikan karakter berdasarkan Kemendikbud di Indonesia, termasuk SMA Methodist-8 Medan. Permasalahan yang muncul

dilapangan adalah belum ada metode yang mampu menentukan karakter peserta didik sehingga pihak sekolah khususnya guru belum bisa memahami secara tepat terhadap karakter yang dimiliki peserta didik. Kurangnya pemahaman karakter peserta didik menjadikan visi misi sekolah belum terlihat tercapai sehingga pendidikan karakter di SMA Methodist-8 Medan belum tepat sasaran.

Dari permasalahan tersebut maka penulis ingin mengangkat judul penelitian yaitu Identifikasi karakter peserta didik menggunakan metode K-Means ( studi kasus SMA Methodist-8 Medan).

## TINJAUAN PUSTAKA

Knowledge Discovery In Database (KDD).

KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining bisa dipakai untuk memperbaiki pengambilan keputusan dimasa depan (Siska, 2016).

### B. Data Mining

Data mining adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting. Data mining biasa juga disebut dengan “Data atau knowledge discovery” atau menemukan pola tersembunyi pada data (Juni, Indrawan dan Dantes, 2016).

### C. K-Means Clustering.

Langkah-langkah melakukan Clustering dengan metode K Means adalah sebagai berikut : (Ong, 2013) :

1. Pilih jumlah cluster K
2. Inisialisasi K pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster dibernilai awal dengan angka-angka random.
3. Alokasi semua data/objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap Identifikasi Karakter Peserta didik. pusat cluster. Jarak antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

$$d_{(i,j)} = \sqrt{(x_{1i} - y_{1j})^2 + (x_{2i} - y_{2j})^2 + \dots + (x_{ki} - y_{kj})^2}$$

Dimana:

- $d_{(x,y)}$  = jarak data ke i ke pusat cluster j  
 $X_{ki}$  = Data ke i pada atribut data ke k  
 $X_{kj}$  = Titik pusat ke j pada atribut ke k

4. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/ objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.
5. Tugaskan lagi tiap objek memakai pusat cluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat cluster tidak

berubah lagi

## METODE PENELITIAN

### A. Pengumpulan Data

Pengumpulan data merupakan tahapan untuk mendapatkan data yang digunakan dalam melakukan penelitian setelah melakukan analisa kebutuhan data. Pengumpulan data dilakukan pada SMA Methodist-8 Medan dikelas 10, 11, dan 12. Adapun atribut karakter yang di gunakan yaitu Jujur, disiplin,percaya diri, peduli,kreatif dan bertanggung jawab kerja keras, mandiri.

Pada saat melakukan tahap penganalisaan data, data di peroleh dari database sekolah SMA Methodist-8 Medan. Setelah data dikumpulkan kemudian data di susun dan di kelompokkan dalam bentuk tabel yang mudah dipahami dan dimengerti dalam tahap pemrosesan dan analisa data. Data yang digunakan dalam penelitian ini adalah data karakter pada siswa yang bersumber dari SMA Methodist-8 Medan. Tahap analisa yaitu proses pembahasan persoalan yang dilakukan sebelum membuat penyelesaian yang akan menjadi hasil utama. Pada bagian analisa ini akan lebih membahas tentang kebutuhan data dan proses clustering dataset dengan menggunakan metode K-Means dalam pengelompokan karakter siswa di SMA Methodist-8 Medan.

### B. KDD (Knowledge Discovery in Database )

Peran KDD pada penggunaan data adalah Menyeleksi Data yang akan digunakan. Terdapat banyak data yang terdapat dalam database, namun hanya beberapa yang dijadikan sebagai data utama untuk melakukan proses. Selanjutnya, data yang sudah ditentukan akan di Cleaning atau dibersihkan serta transformasi data dari beberapa kesalahan-kesalahan sebelum data digunakan.

### C. K-Means Clusterin

Data mining yang digunakan untuk menyelesaikan kasus pengelompokan karakter peserta didik di SMA Methodist-8 Medan pada penelitian ini menggunakan teknik K-Means.

Tabel 1. Data Sampel 1-10 Data Karakter Peserta Didik

Kelas	Jenis Kelamin	Agama	Jujur	Disiplin	Percaya Diri	Peduli	Kreatif	Bertanggung Jawab	Kerja Keras	Mandiri
X-SMA	L	Kristen	1.000	1.000	0.750	0.750	0.750	1.000	1.000	1.000
X-SMA	L	Kristen	0.500	1.000	1.000	0.750	1.000	1.000	1.000	1.000
X-SMA	L	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	1.000	0.750
X-SMA	P	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	0.665	0.750
X-SMA	P	Kristen	1.000	1.000	1.000	1.000	0.665	1.000	1.000	1.000
X-SMA	P	Kristen	1.000	1.000	1.000	1.000	0.665	1.000	0.665	1.000
X-SMA	P	Kristen	0.750	1.000	1.000	0.750	0.750	0.000	1.000	1.000
X-SMA	P	Kristen	0.500	0.750	1.000	1.000	1.000	0.750	1.000	1.000
X-SMA	L	Kristen	1.000	1.000	0.500	0.750	1.000	1.000	1.000	1.000
X-SMA	L	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	1.000	1.000

Langkah selanjutnya adalah menentukan jumlah cluster, adapun cluster yang diinginkan terdiri atas 3 cluster. Kemudian pemberian nilai centroid yang diambil secara acak dari tabel di atas, dengan menggunakan formula :

$$C_i = ( \text{rand}(0,1), \text{rand} (0,1), \text{rand} (0,1))$$

Dimana  $j$  adalah nomor index kriteria atau  $\{j | I_{j_3}, j \in A\}$  dan  $j = 1 \dots 3$ . Langkah berikutnya adalah menghitung jarak menggunakan rumus Euclidean index random centroid nya untuk centroid iterasi 1 adalah sebagai berikut yang mana untuk C1 diambil dari data ke-2 dari 10 data yang akan dilakukan eksperimen, C2 diambil data ke-4 dan C3 diambil data ke-8 .

Tabel 2. Centroid Awal Iterasi (1)

Centroid Awal ( Iterasi 1)	Jujur	Disiplin	Percaya Diri	Peduli	Kreatif	Bertanggung Jawab	Kerja Keras	Mandiri
Data Ke-2 sebagai cluster 1	0.500	1.000	1.000	0.750	1.000	1.000	1.000	1.000
Data Ke-4 sebagai cluster 2	1.000	1.000	0.750	1.000	1.000	1.000	0.665	0.750
Data Ke-8 sebagai cluster 3	0.500	0.750	1.000	1.000	1.000	0.750	1.000	1.000

Data pada tabel 2 digunakan untuk menghitung jarak terdekat centroid C1, C2 dan C3. Menghitung jarak setiap data yang ada terhadap nilai centroid akan dijelaskan sebagai berikut untuk data iterasi pertama dengan rumus :

$$DL_i(x_2, x_1) = \|x_2 - x_1\| \sqrt{\sum_{j=1}^p (x_{2j} - x_{1j})^2}$$

$$DL_i(x_2, x_1) = \|x_2 - x_1\| \sum_{j=1}^p |x_{2j} - x_{1j}|$$

$$X1 = (1.00, 1.000, 0.750, 0.750, 0.750, 1.000, 1.000, 1.000),$$

$$D.1.1 = \sqrt{(1 - 0.5)^2 + (1 - 1)^2 + (0.75 - 1)^2 + (0.75 - 0.75)^2 + (0.75 - 1)^2 + (1 - 1)^2 + (1 - 1)^2}$$

$$C1.1 = (0.500, 1.000, 1.000, 0.750, 1.000, 1.000, 1.000, 1.000),$$

$$= 0.612$$

$$X1 = (1.000, 1.000, 0.750, 0.750, 0.750, 1.000, 1.000, 1.000),$$

$$C2.1 = (1.000, 1.000, 0.750, 1.000, 1.000, 1.000, 0.665, 0.750),$$

$$D2.1 = \sqrt{(1 - 1)^2 + (1 - 1)^2 + (0.75 - 0.75)^2 + (0.75 - 1)^2 + (0.75 - 1)^2 + (1 - 1)^2 + (1 - 0.665)^2 + (1 - 0.75)^2}$$

$$= 0.506$$

Begitu seterusnya sampai data ke-10 hingga diperoleh data sebagai berikut:

Tabel 3. Jarak Euclidean Iterasi 1

Jenis Kelamin	Agama	Jujur	Disiplin	Percaya diri	Peduli	Kreatif	Bertanggung jawab	Kerja Keras	Mandiri	C1	C2	C3	Jarak minimum
L	Kristen	1.000	1.000	0.750	0.750	0.750	1.000	1.000	1.000	0.612	0.550	0.750	0.550
L	Kristen	0.500	1.000	1.000	0.750	1.000	1.000	1.000	1.000	0.000	0.761	0.433	0.000
L	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	1.000	0.750	0.675	0.335	0.724	0.335

Jenis Kelamin	Agama	Jujur	Disiplin	Percaya diri	Peduli	Kreatif	Bertanggung jawab	Kerja Keras	Mandiri	C1	C2	C3	Jarak minimum
P	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	0.665	0.750	0.761	0.000	0.804	0.000
P	Kristen	1.000	1.000	1.000	1.000	0.665	1.000	1.000	1.000	0.652	0.598	0.698	0.598
P	Kristen	1.000	1.000	1.000	1.000	0.665	1.000	0.665	1.000	0.733	0.481	0.774	0.481
P	Kristen	0.750	1.000	1.000	0.750	0.750	0.000	1.000	1.000	1.061	1.230	0.901	0.901
P	Kristen	1.000	1.000	1.000	1.000	1.000	0.750	1.000	1.000	0.612	0.550	0.559	0.550
L	Kristen	1.000	1.000	0.500	0.750	1.000	1.000	1.000	1.000	0.707	0.550	0.829	0.550
L	Kristen	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.559	0.481	0.612	0.481

Langkah berikutnya adalah menentukan letak cluster dengan cara membandingkan antara ke tiga cluster, nilai minimum merupakan nilai yang menjadi pilihan, jika telah ditemukan nilai yang paling kecil (minimum) maka dapat dimasukkan kedalam cluster tersebut. Lebih jelas lihat tabel 4 dibawah ini:

Tabel 4. Posisi Clsuter iterasi 1

Kelas	Jenis Kelamin	Agama	Jujur	Disiplin	Percaya diri	Peduli	Kreatif	Bertanggung jawab	Kerja Keras	Mandiri	C1	C2	C3
X-SMA	L	Kristen	1.000	1.000	0.750	0.750	0.750	1.000	1.000	1.000		●	
X-SMA	L	Kristen	0.500	1.000	1.000	0.750	1.000	1.000	1.000	1.000	●		
X-SMA	L	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	1.000	0.750		●	
X-SMA	P	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	0.665	0.750		●	
X-SMA	P	Kristen	1.000	1.000	1.000	1.000	0.665	1.000	1.000	1.000		●	
X-SMA	P	Kristen	1.000	1.000	1.000	1.000	0.665	1.000	0.665	1.000		●	
X-SMA	P	Kristen	0.750	1.000	1.000	0.750	0.750	0.000	1.000	1.000			●
X-SMA	P	Kristen	1.000	1.000	1.000	1.000	1.000	0.750	1.000	1.000		●	
X-SMA	L	Kristen	1.000	1.000	0.500	0.750	1.000	1.000	1.000	1.000		●	
X-SMA	L	Kristen	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000		●	

Langkah selanjutnya untuk menentukan nilai centroid baru untuk iterasi berikutnya atau iterasi 2, nilai ini ditentukan oleh data yang masuk ke dalam cluster, berdasarkan tabel diatas (data 1-10) diperoleh nilai sebagai berikut:

- o Cluster 1 terdapat 1 data
- o Cluster 2 terdapat 8 data

o Cluster 3 terdapat 1 data

Untuk menentukan nilai centroid baru (contoh pada Cluster 1 terdapat 1 data dapat dilakukan dengan cara berikut: = Untuk centroid kedua (C1) terdapat 1 data:

$$ck = \frac{\text{Jumlah dari nilai yang masuk kedalam cluster}}{\text{Jumlah data yang masuk}}$$

Untuk centroid kedua (C1) terdapat 1 data, C2 terdapat 8 data dan C3 terdapat 1 data dapat dilihat dalam tabel 5 :

Tabel 5. Perhitungan Centroid Kedua

C1	C2	C3
$ck_1 = \frac{0.5}{1} = 0.5$	$ck_1 = \frac{1 + 1 + 1 + 1 + 1 + 1 + 1 + 1}{8} = 1$	$ck_1 = \frac{0.750}{1} = 0.750$
$ck_2 = \frac{1}{1} = 1$	$ck_2 = \frac{1 + 1 + 1 + 1 + 1 + 1 + 1 + 1}{8} = 1$	$ck_2 = \frac{1}{1} = 1$
$ck_3 = \frac{1}{1} = 1$	$ck_3 = \frac{0.750 + 0.750 + 0.750 + 1 + 1 + 1 + 0.500 + 1}{8} = 0.844$	$ck_3 = \frac{1}{1} = 1$
$ck_4 = \frac{0.750}{1} = 0.750$	$ck_4 = \frac{0.750 + 1 + 1 + 1 + 1 + 1 + 0.750 + 1}{8} = 0.938$	$ck_4 = \frac{0.750}{1} = 0.750$
$ck_5 = \frac{1}{1} = 1$	$ck_5 = \frac{0.750 + 1 + 1 + 0.665 + 0.665 + 1 + 1 + 1}{8} = 0.885$	$ck_5 = \frac{0.750}{1} = 0.750$
$ck_6 = \frac{1}{1} = 1$	$ck_6 = \frac{1 + 1 + 1 + 1 + 1 + 0.750 + 1 + 1}{8} = 0.969$	$ck_6 = \frac{0}{1} = 0$
$ck_7 = \frac{1}{1} = 1$	$ck_7 = \frac{1 + 1 + 0.665 + 1 + 0.665 + 1 + 1 + 1}{8} = 0.916$	$ck_7 = \frac{1}{1} = 1$
$ck_8 = \frac{1}{1} = 1$	$ck_8 = \frac{1 + 0.750 + 0.750 + 1 + 1 + 1 + 1 + 1}{8} = 0.938$	$ck_8 = \frac{1}{1} = 1$

Secara keseluruhan diperoleh nilai centroid untuk C1, C2 dan C3 adalah :

Tabel 6. Centroid Awal Iterasi 2

Centroid Awal ( Iterasi 2)	Jujur	Disiplin	Percaya Diri	Peduli	Kreatif	Bertanggung Jawab	Kerja Keras	Mandiri
	0.500	1.000	1.000	0.750	1.000	1.000	1.000	1.000
	1.000	1.000	0.844	0.938	0.885	0.969	0.916	0.938
	0.750	1.000	1.000	0.750	0.750	0.000	1.000	1.000

Untuk mencari nilai centroid iterasi selanjutnya, ulangi langkah seperti contoh iterasi pertama, setelah nilai centroid baru ditemukan maka ulangi langkah perhitungan jarak yaitu pada langkah sebelumnya dan kemudian mencari nilai minimum.

Data akhir adalah jika langkah-langkah di atas diulang dengan langkah-langkah yang sama hingga menghasilkan nilai pada suatu cluster tepat sama dari data sebelumnya atau dengan kata lain tidak ada berubah lagi posisinya pada cluster maka perhitungan centroid berhenti pada iterasi ini yaitu iterasi 2. Posisi pada tiap cluster terakhir pada iterasi kedua dapat dilihat pada table 7 berikut:

Tabel 7. Hasil Akhir Posisi Cluster Iterasi 2

Jenis Kelamin	Agama	Jujur	Disiplin	Percaya diri	Peduli	Kreatif	Bertanggung jawab	Kerja Keras	Mandiri	C1	C2	C3
L	Kristen	1.000	1.000	0.750	0.750	0.750	1.000	1.000	1.000		●	
L	Kristen	0.500	1.000	1.000	0.750	1.000	1.000	1.000	1.000	●		
L	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	1.000	0.750		●	
P	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	0.665	0.750		●	
P	Kristen	1.000	1.000	1.000	1.000	0.665	1.000	1.000	1.000		●	
P	Kristen	1.000	1.000	1.000	1.000	0.665	1.000	0.665	1.000		●	
P	Kristen	0.750	1.000	1.000	0.750	0.750	0.000	1.000	1.000			●
P	Kristen	1.000	1.000	1.000	1.000	1.000	0.750	1.000	1.000		●	
L	Kristen	1.000	1.000	0.500	0.750	1.000	1.000	1.000	1.000		●	
L	Kristen	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000		●	

Berdasarkan data sampel yang diuji diperoleh hasil sebagai berikut :

Tabel 8. Hasil Centroid Akhir iterasi 2

Centroid Awal ( Iterasi 2)	Jujur	Disiplin	Percaya Diri	Peduli	Kreatif	Bertanggung Jawab	Kerja Keras	Mandiri
	0.500	1.000	1.000	0.750	1.000	1.000	1.000	1.000
	1.000	1.000	0.844	0.938	0.885	0.969	0.916	0.938
	0.750	1.000	1.000	0.750	0.750	0.000	1.000	1.000

Dari hasil pengujian data sampel yang dihitung manual menggunakan Algoritma K-Means dapat disimpulkan dari pengelompokan atribut karakternya dan jumlah rata-rata dari setiap cluster dapat disimpulkan anggota cluster 1 memiliki nilai rata-rata centroid 0.637 , cluster 2 dengan nilai rata-rata lebih rendah berjumlah 0,408 dan cluster 3 mempunyai nilai rata-rata tertinggi yaitu 0.984. Cluster 1 secara keseluruhan memiliki nilai karakter pada semua atribut karakter dengan nilai centroid diatas 0,91 yaitu Disiplin (1,000) Percaya diri (1,000), BertanggungJawab (1,000), Kerja Keras (1,000), Mandiri ( 1,000). Cluster 2 memiliki nilai centroid diatas 0,94 yaitu Jujur (1,000), Disiplin (1,000), Peduli (0,944), dan Tanggung Jawab (0,969). Cluster 3 memiliki nilai centroid diatas 0.78 yaitu Disiplin (1,000), Kerja Keras (1,000), dan Mandiri (1,000).

Tabel 9. Jumlah Siswa Pada Setiap Cluster

Cluster	Jujur	Disiplin	Percaya Diri	Peduli	Kreatif	Bertanggung Jawab	Kerja Keras	Mandiri
C1 ( Karakter Baik )	1							

Cluster	Jujur	Disiplin	Percaya Diri	Peduli	Kreatif	Bertanggung Jawab	Kerja Keras	Mandiri
C2 ( Karakter Sangat Baik)	8							
C3 ( Karakter Cukup)	1							

## HASIL DAN PEMBAHASAN

### Implementasi Dan Pengujian

#### A. Implementasi

Pengujian terhadap analisa sangat penting di lakukan untuk menemukan dan memastikan apakah hasil dari analisa secara manual menggunakan data mining algoritma K-Means sesuai dengan keputusan yang di harapkan dengan pengujian data menggunakan software Bahasa R.

#### B. Pengujian

Tujuan pengujian perangkat lunak ini adalah untuk membuktikan bahwa pengelompokan karakter Peserta Didik pada SMA Methodist-8 Medan sesuai dengan perhitungan yang menggunakan Microsoft Excel berdasarkan rumus algoritma K-Means. Pada penelitian pengujian menggunakan Bahasa R yang telah banyak digunakan untuk pengujian Data Mining. Adapun data awal yang digunakan adalah sebagai berikut:

Tabel 10. Data Awal 10 Data (Data Sampel)

Kelas	Jenis Kelamin	Agama	Jujur	Disiplin	Percaya Diri	Peduli	Kreatif	Bertanggung Jawab	Kerja Keras	Mandiri
X-SMA	L	Kristen	1.000	1.000	0.750	0.750	0.750	1.000	1.000	1.000
X-SMA	L	Kristen	0.500	1.000	1.000	0.750	1.000	1.000	1.000	1.000
X-SMA	L	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	1.000	0.750
X-SMA	P	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	0.665	0.750
X-SMA	P	Kristen	1.000	1.000	1.000	1.000	0.665	1.000	1.000	1.000
X-SMA	P	Kristen	1.000	1.000	1.000	1.000	0.665	1.000	0.665	1.000
X-SMA	P	Kristen	0.750	1.000	1.000	0.750	0.750	0.000	1.000	1.000
X-SMA	P	Kristen	0.500	0.750	1.000	1.000	1.000	0.750	1.000	1.000
X-SMA	L	Kristen	1.000	1.000	0.500	0.750	1.000	1.000	1.000	1.000
X-SMA	L	Kristen	1.000	1.000	0.750	1.000	1.000	1.000	1.000	1.000

Dari hasil pengujian menggunakan 10 data sampel karakter peserta didik dengan Bahasa R dan manual dapat disimpulkan bahwa siswa yang memiliki karakter Cukup (C1) 1 siswa , untuk karakter Baik (C2) ada 8 siswa dan untuk Karakter Sangat Baik (C3) 1 Siswa . Berikut adalah Nilai Centroid perhitungan manual iterasi 2 yang merupakan iterasi terakhir telah sama dengan centroid yang dihitung menggunakan Bahasa R.

Tabel 11. Nilai Centroid Data Sampel

Jujur	Disiplin	Percaya Diri	Peduli	Kreatif	Tanggung Jawab	Kerja Keras	Mandiri
0.50	1.000	1.000	0.750	1.000	1.000	1.000	1.000
1.00	1.000	0.844	0.938	0.885	0.916	0.916	0.938
0.75	1.000	1.000	0.750	0.750	1.000	1.000	1.000

Berdasarkan pengujian dengan data sampel menggunakan Bahasa R maka diperoleh pengelompokan pada cluster 1 terdapat 1 data siswa yang berkarakter Cukup , 8 data siswa yang berkarakter Baik dan 1 data siswa berkarakter sangat baik. Berikut ini adalah hasil perhitungan seluruh data yaitu 230 data siswa SMA Methodist-8 Medan menggunakan Bahasa R:

Tabel 12. Hasil Pengujian Seluruh Data

Jujur	Disiplin	Percaya Diri	Peduli	Kreatif	Tanggung Jawab	Kerja Keras	Mandiri
0.625	1.000	0.545	0.800	0.949	0.545	1.000	1.000
1.000	1.000	0.920	0.800	0.595	0.595	0.888	1.000
0.996	1.000	0.916	0.800	0.409	0.407	1.934	0.901

Dari tabel di atas didapatkan hasil sebagai berikut:

1. Cluster 1 yang merupakan cluster yang memiliki nilai rata-rata centroid tertinggi yaitu 0,808. Karakter Disiplin yang berada pada urutan pertama nilai centroid tertinggi yaitu (1,000), urutan kedua adalah karakter Kerja Keras (1,000) dan kemudian karakter Mandiri (1,000), dan karakter Kreatif (0,949).
2. Cluster 2 yang merupakan cluster yang memiliki nilai rata-rata centroid tertinggi yaitu 0,850. Karakter Jujur yang berada pada urutan pertama nilai centroid tertinggi yaitu (1,000), urutan kedua adalah karakter disiplin (1,000) dan kemudian karakter dan mandiri (1,000), kemudian karakter Percaya Diri (0,920), dan Karakter Kerja Keras (0,888).
3. Cluster 3 memiliki nilai rata-rata centroid yang lebih rendah 0,795. Karakter yang Cukup pada cluster ini adalah karakter Bertanggung Jawab (0,407) dan kreatif (0,409) yang memiliki nilai centroid dibawah nilai rata-rata.

Untuk mengetahui lebih rinci posisi cluster karakter siswa dari kelas X-SMA sampai kelas XII-SMA adalah sebagai berikut:

Tabel 13. Cluster Jumlah Siswa Seluruh Data

Cluster	Kelas X-SMA	Kelas XI-SMA	Kelas XII-SMA
C1	16 Siswa	16 Siswa	16 Siswa
C2	24 Siswa	20 Siswa	24 Siswa
C3	40 Siswa	34 Siswa	40 Siswa

Tabel diatas menunjukkan jumlah pengelompokan karakter berdasarkan kelas yaitu kelas X yang memiliki karakter Sangat Baik (C1) 16 siswa dari 80 siswa, kelas XI berjumlah 16 dari 70 siswa, kelas XII berjumlah 16 dari 80 siswa. Sedangkan cluster dengan karakter Baik (C2) kelas X berjumlah 24 siswa, kelas XI berjumlah 20 siswa, kelas XII berjumlah 24 siswa, dan cluster Cukup (C3) Kelas X berjumlah 40 Siswa, kelas XI berjumlah 34 siswa dan kelas XII berjumlah 40 siswa.

## KESIMPULAN

### A. Kesimpulan

Adapun kesimpulan berdasarkan proses clustering menggunakan Bahasa R dapat disimpulkan sebagai berikut:

1. Pengelompokan dengan menggunakan algoritma K-Means menghasilkan pengelompokan karakter yaitu karakter Cukup (C1) , Karakter Baik (C3) dan karakter Sangat Baik (C2) yang ada di SMA Methodist-8 Medan.
2. Karakter yang paling kuat diterapkan siswa disekolah adalah karakter disiplin, kerja keras, mandiri dan kreatif dengan nilai rata-rata (0,850) Sedangkan karakter yang baik diterapkan siswa adalah disiplin, kerja keras, mandiri dan percaya diri dengan nilai rata-rata (0,808) dan karakter Cukup yang masih lemah diterapkan disekolah adalah bertanggungjawab dan kreatif dengan nilai rata-rata ( 0.795).
3. Karakter kuat dapat dikembangkan dengan cara sekolah terus melakukan kebiasaan-kebiasaan yang melibatkan siswa agar karakter dalam diri siswa dapat terlihat sedangkan untuk karakter bertanggungjawab dan kreatif agar tidak lemah maka pihak sekolah selalu memberikan bimbingan kepada siswa dan memberi contoh kebiasaan-kebiasaan baik serta kegiatan yang dapat diikuti siswa disekolah.

### B. Saran

1. Pada penelitian tentu nya tidak terlepas dari kelemahan dan kekurangan. Adapun saran yang peneliti berikan yaitu mengingat pentingnya pendidikan karakter sebagai sarana untuk pengembangan dan membentuk watak kepribadian peserta didik dan peradaban bangsa yang berkarakter diharapkan hasil penelitian ini membantu pihak sekolah untuk menindak lanjuti hal-hal yang perlu dilakukan untuk mengembangkan karakter peserta didik yang belum tercapai sesuai dengan visi dan misi SMA Methodist-8 Medan.
2. Dalam pengembangan karakter bertanggungjawab, sekolah harus lebih memperhatikan kehadiran siswa, membiasakan membuang sampah pada tempatnya, membiasakan hadir tepat waktu datang kesekolah, membiasakan menolong atau membantu orang lain dan membiasakan konsultasi kepada guru kelas atau guru lain sesuai kebutuhan.
3. Dalam mengembangkan karakter kreatif, yang harus dilakukan sekolah adalah membuat kegiatan terprogram yaitu keguatan yang dilakukan bertahap sesuai waktu yang ditetapkan misalnya kegiatan memperingati hari-hari besar nasional, kegiatan kegiatan lomba mata pelajaran seperti matematika, mewarnai, dan lomba karya seni.
4. Dalam pengembangan untuk penelitian selanjutnya yaitu menggabungkan Algoritma K-Means dengan Fuzzy C-Means atau menggabungkan beberapa metode lainnya.

## REFERENSI

- Kasini., Jufriadif Na'am,Sumijan, (2018), "Identifikasi Karakter Siswa Menggunakan Metode K-Means (Studi Kasus SDN 156 Pekanbaru)", Jurnal Inovasi Teknik Informatika, Jiti, Vol.1, No.1.
- Lin Parlina, Agus erdana Windarto, Anjar Wanto, M.Ridwan Lubis, (2018) " 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, Hal 87-93.
- Siska S.T. (2016), "Analisa Dan Penerapan Data Mining Untuk Menentukan Kubikasi Air Terjual

Berdasarkan Pengelompokan Pelanggan Menggunakan Algoritma K-Means Clustering”, Jurnal Teknologi Informasi dan Pendidikan, Vol. 9 No.1 Hal.86-93

Windha Mega Pradnya Duhita, (2015), “Clustering Menggunakan Metode K-Means Untuk Menentukan Status Gizi Balita”, Jurnal Informatika, Vol 15 No. 2.

Prasetyo E (2014), Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab, Andi, Jakarta, Hal.189

Lisa'diyah Ma' rifataini, “ Pengaruh Kegiatan Intrakurikuler dan Ekstrakurikuler terhadap pembentukan Karakter Siswa Sekolah Menengah Atas Negeri (SMAN) 09 Bandar Lampung”, Jurnal P, enelitian Pendidikan Agama dan Keagamaan, Vol 14. Hal 171-186.

Rifkie Primartha, 2018 Belajar Machine Learning Teori dan Praktek. Bandung: Informatika Bandung, Hal. 277