JDSP
Jurnal Data Science Penusa

Implementasi Data Mining Penentuan Kelas Siswa Unggulan Menggunakan Algoritma K-Means Clustering Pada Mts Swasta Pp Al Qomariah

Dedi Nurwanda[1)*], Yudha Perwira[2)]
[1,2)] STMIK Pelita Nusantara, Medan, [1)] Email: dey.al.adlu@gmail.com [2)] Email: yudaperwira25@gmail.com

Received: 12 Agustus 2024 Accepted: 11 Oktober 2024 Published: 27 Desember 2024
Pg. 568-578

***Dedi Nurwanda**

Kata Kunci: K-Means Clustering, Data Mining, Penentuan Kelas, Pendidikan, PHP, MySQL.

JDSP: Jurnal Data Science Penusa is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Abstrak : Penentuan kelas siswa unggulan sering kali dilakukan secara subjektif, sehingga berpotensi tidak optimal dalam memetakan kemampuan akademik siswa. Penelitian ini bertujuan mengimplementasikan algoritma K-Means Clustering dalam pengelompokan siswa di MTS Swasta PP Al Qomariah untuk menentukan kelas unggulan, baik, dan cukup secara objektif. Data yang digunakan berupa nilai tes masuk siswa baru tahun ajaran 2023/2024 dengan lima kriteria: Matematika, Bahasa Inggris, Bahasa Indonesia, Pendidikan Agama Islam, dan Bahasa Arab. Hasil dari pengelompokan ini membagi siswa menjadi tiga cluster, yaitu kelas unggul (C1), baik (C2), dan cukup (C3). Berdasarkan perhitungan K-Means, diketahui hasil olah pengujian Cluster dengan pembagian Cluster Siswa unggul (C1) 6 data siswa, yaitu nomor 1, 5, 8, 12, 13, 17, Cluster siswa baik (C2) 9 data siswa, yaitu 2, 4, 6, 9, 11, 14, 16, 19, 20, Cluster siswa cukup (C3) 5 data siswa, yaitu 3, 7, 10, 15, 18. Penelitian menghasilkam sistem yang dibangun menggunakan bahasa pemrograman PHP dan database MySQL. Dengan penerapan data mining, proses penentuan kelas siswa menjadi lebih efisien dan akurat, memberikan manfaat signifikan bagi pengelolaan pendidikan di sekolah tersebut.

PENDAHULUAN

Perkembangan teknologi komputer yang begitu pesat telah membawa dampak yang sangat besar pada semua aspek kehidupan manusia. Di era teknologi informasi ini, mampu mempermudah segala pekerjaan manusia dalam menjalani kesehariannya. Salah satu aspek yang sangat dibutuhkan pada masa kini adalah bidang pendidikan. Dalam bidang pendidikan, teknologi sangatlah membantu dalam penentuan kelas unggulan sebagai salah satu upaya meningkatkan mutu pendidikan diberbagai sekolah melalui layanan pembelajaran kelas sesuai dengan kriteria siswa.

TINJAUAN LITERATUR

A. Data Mining

Data mining adalah proses ekstraksi informasi yang melekat atau tersembunyi tetapi tidak diketahui dan berguna dari data tidak terstruktur dimana didalamnya terdapat algoritma algoritma untuk memproses data berukuran besar. Data mining istilah yang digunakan untuk menggambarkan penemuan pengetahuan dalam suatu database atau sering disebut Knowledge Discovery in Database (KDD) Konsep data mining dapat diimplementasikan dengan menggunakan banyak algoritma berbeda yang memiliki fungsionalitas untuk menyelesaikan tugas tergantung pada karakteristik data, algoritma data mining harus mengikuti model yang menjelaskan tugas data mining, dimana model data mining dapat diklasifikasikan sebagai model prediktif atau model deskriptif[1][2][3].

B. Algoritma K-Means

Algoritma K-Means dilakukkan untuk mengelompokkan data kedalam bentuk satu atau lebih kelompok, dimana data yang memiliki reprensentative persamaan dikelompokkan dalam satu kelompok dan data yang

memiliki perbedaan masuk kedalam kelompok yang lain[4][5].
Algoritma K-Means merupakan salah satu metode yang ada di dalam data mining yang melakukan pemodelan untuk pegelompokan data dengan sistem partisi. K-Means juga merupakan salah satu algoritma pengelompokan paling popular dan sederhana, dimana algoritma ini dipakai mulai tahun 1955 dan sampai sekarang masih banyak digunakan. Algoritma K-means sangat membantu dalam mengelompokkan data dan juga sangat berpengaruh pada hasil cluster dengan menggunakan data yang besar Langkah-langkah dalam Algoritma K-means Clustering [5]:

1. Menentukan jumlah cluster
2. Menentukan nilai centroid (titik pusat cluster) secara acak.
3. Menghitung jarak antara data atau objek dengan pusat cluster menggunakan teori Euclidian Distance yang dirumuskan sebagai berikut: $D(i,j) = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + \cdots + (Xki - Xkj)^2}$ ...(2.1)
   Dimana :
   D (i,j) : Jarak data ke i kepusat cluster j
   Xki : Data ke i pada atribut data ke k
   Xkj : Titik pusat ke j pada atribut data ke k
4. Data ditempatkan dalam cluster yang terdekat.
5. Pusat cluster baru akan ditentukan bila semua data telah ditetapkan dalam cluster terdekat dengan rumus berikut: D=1/n .....(2.2)
   Dimana D : Centroid/Cluster Baru
   n : Banyaknya data pada cluster
6. Proses penentuan pusat cluster dan penempatan data dalam cluster diulangin sampai nilai centroid tidak berubah lagi.

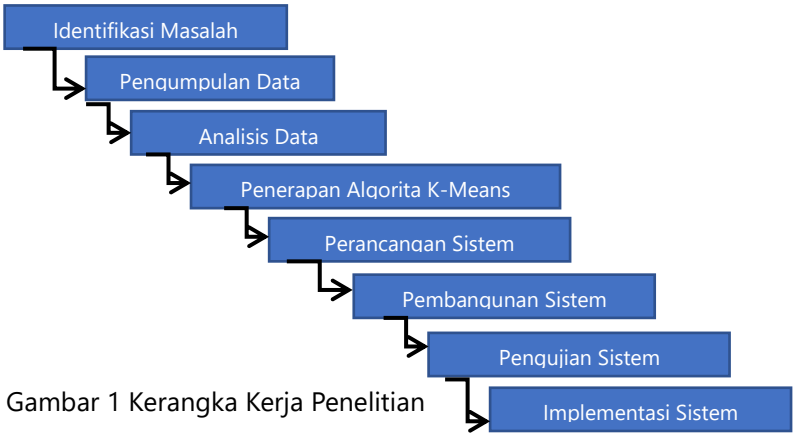
C. Cluster

Clustering adalah proses pemetaan kelompok objek yang sama. Kelompok-kelompok yang dibentuk oleh pengelompokan disebut sebagai cluster[8]. Clustering merupakan suatu teknik data mining yang membagi-bagikan data ke dalam beberapa kelompok (grup atau cluster atau segmen) sesuai dengan yang diinginkan, dimana setiap cluster atau kelompok dapat ditempati beberapa anggota secara bersamaan dan di dalam kelompok bisa disebutkan objek atau variabel yang sama[5]

METODE PENELITIAN

Kerangka Kerja Penelitian

Kerangka kerja penelitian merupakan tahap yang dilakukan dalam penelitian yang bertujuan untuk mencapai hasil yang maksimal. Suatu konsep didalam penelitian yang menghubungkan antara satu dengan lainnya, sehingga penelitian menjadi tersusun secara sistematis dan dapat diterima oleh semua pihak, berikut merupakan kerangka kerja penelitian yang dilakukan oleh peneliti dalam menyelesaikan permasalahan:



Gambar 1 Kerangka Kerja Penelitian

A. Uraian Kerangka Penelitian

Berdasarkan kerangka kerja penelitian yang telah dibuat di atas, maka dapat diuraikan pembahasan masing-masing tahap dalam penelitian sebagai berikut;

1. Mengidentifikasi Masalah
   Langkah Pertama dalam Penelitian ini mengidentifikasi masalah untuk mengetahui permasalahan apa yang ada dalam pengelompokan penentuan kelas siswa unggulan kepada MTS Swasta PP Al Qomariah, dengan demikian peneliti lebih memahami permasalahan yang akan diteliti. Dari observasi, wawancara dan pengambilan sampel nilai tes masuk Penerimaan Peserta Didik Baru (PPDB) yang dilakukan oleh peneliti sejauh ini, belum adanya penerapan algoritma dalam pengelompokan kelas siswa unggulan dalam hal ini MTS Swasta PP Al Qomariah.
2. Pengumpulan Data
   Peneliti mengumpulkan data dengan beberapa kegiatan yaitu melakukan wawancara, observasi, kuisioner dan mengumpulkan literatur yang berhubungan dengan pengelompokan kelancaran pembayaran piutang pada mitra.
   a. Observasi
      Peneliti melakukan pengamatan terhadap proses pembagian kelas pada penerimaan peserta didik baru yang dilakukan untuk memahami mekanisme dan kriteria yang digunakan dalam mengelompokkan siswa. Pengamatan ini melibatkan analisis terhadap prosedur yang diterapkan oleh pihak madrasah, mulai dari pengumpulan Data Primer seperti nilai ujian masuk peserta didik baru, hingga pembagian kelas kedalam kelas Unggul, kelas Baik, dan kelas Cukup yang mempertimbangkan berbagai aspek, termasuk kesetaraan jumlah siswa per kelas dan keseimbangan kemampuan akademik
   b. Wawancara
      Wawancara dilakukan dengan Bapak Mhd Habib Aulia, S.Pd., selaku Kepala Madrasah untuk mengidentifikasi mekanisme yang diterapkan dalam pembagian kelas pada penerimaan peserta didik baru, serta untuk memperoleh perspektif tentang efektivitas mekanisme tersebut dalam menyesuaikan penempatan siswa berdasarkan kemampuan akademik dan kebutuhan pembelajaran mereka.
   c. Studi Literatur
      Peneliti mengumpulkan Data Sekunder untuk kebutuhan Landasan Teori dan Anaslisis Data dalam penerapan algoritma K-Means Clustering penentuan kelas siswa unggulan dari sumber tepercaya yaitu dari Buku atau Jurnal Ilmiah atau Internasional yang terbit mulai dari Tahun 2019 s.d 2024.

3. Analisa Data
   Pada Tahapan ini dilakukan proses memilah dan merangkum data – data yang layak digunakan sebagai bahan dari penerapan algoritma K-Means Clustering, mulai dari Data Pendaftar Peserta Didik Baru, Data Nilai Tes Penerimaan Peserta Didik Baru dan masing-masing kriteria yang digunakan.
4. Penerapan Algoritma *K-Means Clastering*
   Pada tahap ini akan dilakukan penerapan algoritma K-Means Clustering dalam melakukan perhitungan untuk mendapatkan hasil data pengelompokan ke dalam 3 Cluster yaitu Kelas Unggul, Kelas Baik, dan Kelas Cukup, dalam hal ini MTS PP Al Qomariah.
5. Perancangan Sistem
   Dalam tahap ini sistem akan dirancang menggunakan perancangan UML (Unifield Modeling Language). UML menyediakan berbagai diagram yang memungkinkan perancangan sistem dilakukan secara komprehensif, baik dari sisi fungsionalitas, struktur, maupun perilaku sistem. Dengan menggunakan UML, komponen-komponen utama dari sistem, seperti aktor, entitas, serta alur proses, dapat divisualisasikan dalam bentuk diagram seperti use case diagram, class diagram, activity diagram, dan sequence diagram.
6. Pembagunan Sistem
   Dalam tahap ini system akan dibangun dengan berbasis Web dengan menggunakan bahasa pemrograman Visual Studio dan MySQL sebagai Database Management System (DBMS).
7. Pengujian Sistem
   Pada tahap ini sistem yang sudah dibangun akan diuji coba sebelum diimplementasikan. Metode yang digunakan peneliti dalam pengujian sitem yaitu menggunakan Black Box testing. Hasil dari pengujian

tersebut akan memberikan umpan balik mengenai stabilitas dan kehandalan sistem. Temuan berupa bug, kesalahan logika, atau masalah performa dapat segera ditindaklanjuti sebelum sistem dirilis ke pengguna akhir.

8. Implementasi Sistem
   Pada tahap ini dilakukan implementasi sistem yang sudah selesai dirancang dan dibangun dan setelah diperbaiki kesalahan setelah pengujian sistem. Dan untuk sekarang ini sistem yang sudah dibuat oleh peneliti masih digunakan untuk penyelesaian Tugas Akhir untuk mengambil gelar strata – 1, dan kedepannya Peneliti berharap Sistem ini juga akan berkelanjutan digunakan oleh Lembaga yang sudah peneliti lakukan Riset yang akan kemudian peneliti akan melakukan pemantau dan penmeliharaan terhadap sistem secara berkelanjutan dimasa yang akan datang.

B. Pemrosesan Dataset Nilai Siswa

Penenlitian ini adalah mengelompokkan kelas siswa ke dalam 3 cluster yaitu kelas siswa unggul (C1), kelas siswa baik (C2) dan kelas siswa cukup (C3). Dataset yang disiapkan sebagai sampel adalah sebanyak 20 data yang sudah ditrasformasikan ke bentuk angka yang sudah siab diolah algoritma k-mens sebagai algoritma yang diterapkan pada penelitian ini. Adapun datasetnya adalah sebagai berikut:

Tabel 4.1 Data Nilai Siswa

No	Nama Siswa	Nilai Per Bidang Studi				
		B. ING	B. IND	MM	PAI	Bhs Arab
1	MUHAMMAD ATTAR IBNU HAYAN	75	93	70	88	90
2	MUHAMMAD KHAIRUL AZMAN	70	100	60	77	90
3	MUHAMMAD TAVAN RAHMAN SIPAYUNG	70	93	70	60	90
4	MUHAMMAD WILDAN HASIBUAN	70	100	60	83	90
5	RADITYA SABTA ARKINDRA	75	93	60	88	90
6	RIFA'I NURIL SYAHBI	70	100	60	77	90
7	KHALIL AZHARI	70	93	60	60	90
8	M.FADLI GUNAWAN	75	93	60	88	90
9	HARUN SAH ARRASYD	70	100	60	77	90
10	ZASKIA KIRANA	70	93	70	60	90
11	PUTRI AYU	70	100	60	83	90
12	RAYSA SAVIRA	75	100	60	88	90
13	RIZKY RIVANI TANJUNG	75	93	60	88	90
14	SEFIRA AMELIA	70	100	60	77	90
15	SOFIYATUL HUSNA	70	93	70	60	90
16	SULISTIA ANGGRIANI	70	100	60	83	90
17	SYAFRIDA HANDAYANI	75	100	60	88	90
18	SYAHIDA AZAHRI YANTI	70	87	70	72	90
19	SYIFA AQILLA	75	100	60	70	90
20	RAFA FAHREZA	70	100	60	84	90

Dari data yang ada pada tabel 4.1 diatas akan dikelompokkan ke dalam 3 cluster yaitu kelas siswa unggul (C1), kelas siswa baik (C2) dan kelas siswa cukup (C3) menngunakan algoritma K-Means Clustering, dimana hasil pengelompokannya adalah pada tabel berikut:

Tabel 4.2 Hasil Cluster Data Kelas Siswa

No	Kode	B. ING	B. IND	MM	PAI	B. ARAB	C1	C2	C3	Jarak Terdekat	Group
1	A001	75	93	60	88	90	2	12	27	2	C1
2	A002	70	100	60	77	90	13	2	18	2	C2
3	A003	70	93	70	60	90	30	22	3	3	C3
4	A004	70	100	60	83	90	8	4	24	4	C2
5	A005	75	93	60	88	90	2	12	27	2	C1
6	A006	70	100	60	77	90	13	2	18	2	C2
7	A007	70	93	60	60	90	29	20	8	8	C3
8	A008	75	93	60	88	90	2	12	27	2	C1
9	A009	70	100	60	77	90	13	2	18	2	C2
10	A010	70	93	70	60	90	30	22	3	3	C3
11	A011	70	100	60	83	90	8	4	24	4	C2
12	A012	75	100	60	88	90	4	10	28	4	C1
13	A013	75	93	60	88	90	2	12	27	2	C1
14	A014	70	100	60	77	90	13	2	18	2	C2
15	A015	70	93	70	60	90	30	22	3	3	C3
16	A016	70	100	60	83	90	8	4	24	4	C2
17	A017	75	100	60	88	90	4	10	28	4	C1
18	A018	70	87	70	72	90	21	18	11	11	C3
19	A019	75	100	60	70	90	19	10	15	10	C2
20	A020	70	100	60	84	90	8	5	24	5	C2

Dari hasil yang ada pada tabel 4.2 diatas maka, Kelas Siswa Unggul (C1) sebanyak 6 data siswa, yaitu nomor (1, 5, 8, 12, 13, 17), kelas siswa baik (C2) sebanyak 9 data siswa, yaitu (2, 4, 6, 9, 11, 14, 16, 19, 20), sedangkan kelas siswa cukup (C3) sebanyak 5 data siswa, yaitu (3, 7, 10, 15, 18).

A. Penerapan Algoritma K-Means Clustering

1. Input Data dan Nilai Attribut Siswa

   Pada sesi ini akan ditampilkan data berupah nilai setiap setiap attribut siswa, dimana data tersebut akan diolah dengan tahapan algoritma k-means.

Tabel 5.1 Data Nilai Attribut siswa

No	Kode Alternatif	B. ING (X)	B. IND (y)	MM (z)	PAI (aa)	B. ARAB (ab)
1	A001	75	93	60	88	90
2	A002	70	100	60	77	90
3	A003	70	93	70	60	90

4	A004	70	100	60	83	90
5	A005	75	93	60	88	90
6	A006	70	100	60	77	90
7	A007	70	93	60	60	90
8	A008	75	93	60	88	90
9	A009	70	100	60	77	90
10	A010	70	93	70	60	90
11	A011	70	100	60	83	90
12	A012	75	100	60	88	90
13	A013	75	93	60	88	90
14	A014	70	100	60	77	90
15	A015	70	93	70	60	90
16	A016	70	100	60	83	90
17	A017	75	100	60	88	90
18	A018	70	87	70	72	90
19	A019	75	100	60	70	90
20	A020	70	100	60	84	90

2**.** Menentukan Jumlah Cluster

Data yang ada pada tabel 5.1 diatas akan dikelompokkan ke dalam 3 *cluster* dalam mengelompokkan kelancaran pembayaran piutang oleh nasabah kepada koperasi CKM, yaitu kelas siswa unggul (C1), kelas siswa baik(c2), kelas siswa cukup(C3).

3. Menentukan Nilai Centroid (Titik Pusat Cluster)

Pada tahap ini akan ditentukan pusat *cluster* pertama secara acak seperti yang terlihat pada tabel berikut.

Tabel 5.2 Tabel Pusat Cluster

Data Ke-	Pusat Cluster	B. ING (X)	B. IND (y)	MM (z)	PAI (aa)	B. ARAB (ab)
1	C1	75	93	70	88	90
2	C2	70	100	60	77	90
3	C3	70	93	70	60	90

4. Menghitung Nilai Jarak

Akan dilakunan perhitungan antara data pertama yang ditabel 5.1 ke pusat *cluster* "centroid" yang ada pada tabel 5,2 diatas sebagai iterasi pertama. Berikut ini merupakan contoh perhitungan :

$$D(i,j) = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + \cdots + (Xki - Xkj)^2} \quad \text{.....(1)}$$

Dimana, D (i,j) : Jarak data ke i kepusat cluster j

Xki : Data

Xkj : Titik pusat cluster

Perhitungan data pertama ke pusat *Cluster* Sebagai berikut:

C1=$\sqrt{(75-75)^2 + (93-93)^2 + (60-60)^2 + (88-88)^2 + (90-90)^2} = 0$

C2=$\sqrt{(75-70)^2 + (93-100)^2 + (60-60)^2 + (88-77)^2 + (90-90)^2} = 14$

C3=$\sqrt{(75-70)^2+(93-93)^2+(60-70)^2+(88-60)^2+(90-90)^2}$ = 30

Setelah itu akan dilakukan perhitungan dari data ke-2 sampai data ke-20 ke semua pusat *cluster*.

5. Menempatkan Data Ke Cluster Terdekat

Setelah semua perhitungan dilakukan dari data ke-1 sampai data ke-20 ke semua pusat *cluster* , maka hasilnya dapat dilihat pada tabel berikut:

Tabel 5.3 Hasil Perhitungan Semua Data Ke Pusat Cluster "Centroid"

No	Kode Alternatif	B. ING	B. IND	MM	PAI	B. ARAB	C1	C2	C3	Jarak Terdekat	Group
1	A001	75	93	70	88	90	0	14	30	0	C1
2	A002	70	100	60	77	90	14	0	21	0	C2
3	A003	70	93	70	60	90	30	21	0	0	C3
4	A004	70	100	60	83	90	10	6	26	6	C2
5	A005	75	93	60	88	90	0	14	30	0	C1
6	A006	70	100	60	77	90	14	0	21	0	C2
7	A007	70	93	60	60	90	28	18	10	10	C3
8	A008	75	93	60	88	90	0	14	30	0	C1
9	A009	70	100	60	77	90	14	0	21	0	C2
10	A010	70	93	70	60	90	30	21	0	0	C3
11	A011	70	100	60	83	90	10	6	26	6	C2
12	A012	75	100	60	88	90	7	12	31	7	C1
13	A013	75	93	60	88	90	0	14	30	0	C1
14	A014	70	100	60	77	90	14	0	21	0	C2
15	A015	70	93	70	60	90	30	21	0	0	C3
16	A016	70	100	60	83	90	10	6	26	6	C2
17	A017	75	100	60	88	90	7	12	31	7	C1
18	A018	70	87	70	72	90	21	17	14	14	C3
19	A019	75	100	60	70	90	19	9	17	9	C2
20	A020	70	100	60	84	90	9	7	27	7	C2

Pada tahap perhitungan iterasi pertama , dapat dilihat dari hasil pengelompokan Kelas Siswa Unggul (C1) sebanyak 6 data siswa, yaitu nomor (1, 5, 8, 12, 13, 17), kelas siswa baik (C2) sebanyak 9 data siswa, yaitu (2, 4, 6, 9, 11, 14, 16, 19, 20), sedangkan kelas siswa cukup (C3) sebanyak 5 data siswa, yaitu (3, 7, 10, 15, 18).

6. Menentukan Pusat Cluster Baru

Langka selanjutnya adalah menentukan pusat cluster baru dengan menjumlahkan semua data setia cluster sesuai rumus yang ada, yakni dengan rumus:

$D = \frac{1}{n}$ .....(2)

Dimana D : Centroid/*Cluster* Baru

1/n : Banyaknya data pada *cluster* dibagi jumlah cluster

Dengan rumus tersebut dilakukan perhitungan untuk menentukan cluster "Centroid" baru.

Untuk Nilai C1 :

C1 = (75+75+75+75+75+75)/6 = 75 Untuk Bhs Inggris

C1 = (93+93+93+100+93+100)/6 = 96 Untuk Bhs Indonesia
C1 = (60+60+60+60+60+60)/6 = 60 Untuk Matematika
C1 = (88+88+88+88+88+88)/6 = 88 Untuk Pend. Agama Islam
C1 = (90+90+90+90+90+90)/6 = 90 Untuk Bhs Arab
Untuk Nilai C2:
C2 = (70+70+70+70+70+70+75+70)/9 = 71 Untuk Bhs Inggris
C2 = (100+100+100+100+100+100+100+100+100)/9 = 100 Untuk Bhs Indonesia
C2 = (60+60+60+60+60+60+60+60+60)/9 = 60 Untuk Matematika
C2 = (77+83+77+77+83+77+83+70+84)/9 = 79 Untuk Pend. Agama Islam
C2 = (90+90+90+90+90+90+90+90+90)/14 = 90 Untuk Bhs Arab
Untuk Nilai C3:
C3 = (70+70+70+70+70)/5 = 70 Untuk Bhs Inggris
C3 = (93+93+93+93+87)/5 = 92 Untuk Bhs Indonesia
C3 = (70+60+70+70+70)/5 = 68 Untuk Matematika
C3 = (60+60+60+60+72)/5 = 62 Untuk Pend. Agama Islam
C3= (90+90+90+90+90)/5 = 90 Untuk Bhs Arab
Maka didapatkan hasil pusat cluster "Centroid" baru untuk iterasi kedua sebagai berikut .

Tabel 5.4 Pusat *Cluster* "Centroid" baru

Data Ke-	Pusat Cluster	B. ING (X)	B. IND (y)	MM (z)	PAI (aa)	B. ARAB (ab)
1	C1	75	96	60	88	90
2	C2	71	100	60	79	90
3	C3	70	92	68	62	90

7. Menghitung Kembali Nilai Jarak Sampai Tidak Berubah
Pada bagian ini akan dilakukan kembali perhitungan antara pusat cluster baru iterasi kedua yang ada di tabel 5.4 ke data nilai siswa pada tabel 5.1 Sesuai Rumus yang ada Setelah dilakukan perhitungan hasilnya tidak berubah dengan hasil yang berada pada tabel tabel 5.3 diatas, sehingga perhitungan mengelompokkan kelas siswa di MTS Swasta PP AL Qomariah di hentikan.

Tabel 5.4 Hasil Perhitungan Semua Data Ke Pusat Cluster "Centroid" baru

No	Kode Alternatif		B. ING	B. IND	MM	PAI	B. ARAB	C1	C2	C3	Jarak Terdekat	Group
1	A001		75	93	60	88	90	2	12	27	2	C1
2	A002		70	100	60	77	90	13	2	18	2	C2
3	A003		70	93	70	60	90	30	22	3	3	C3
4	A004		70	100	60	83	90	8	4	24	4	C2
5	A005		75	93	60	88	90	2	12	27	2	C1
6	A006		70	100	60	77	90	13	2	18	2	C2
7	A007		70	93	60	60	90	29	20	8	8	C3
8	A008		75	93	60	88	90	2	12	27	2	C1
9	A009		70	100	60	77	90	13	2	18	2	C2
10	A010		70	93	70	60	90	30	22	3	3	C3
11	A011		70	100	60	83	90	8	4	24	4	C2
12	A012		75	100	60	88	90	4	10	28	4	C1

13	A013		75	93	60	88	90	2	12	27	2	C1
14	A014		70	100	60	77	90	13	2	18	2	C2
15	A015		70	93	70	60	90	30	22	3	3	C3
16	A016		70	100	60	83	90	8	4	24	4	C2
17	A017		75	100	60	88	90	4	10	28	4	C1
18	A018		70	87	70	72	90	21	18	11	11	C3
19	A019		75	100	60	70	90	19	10	15	10	C2
20	A020		70	100	60	84	90	8	5	24	5	C2

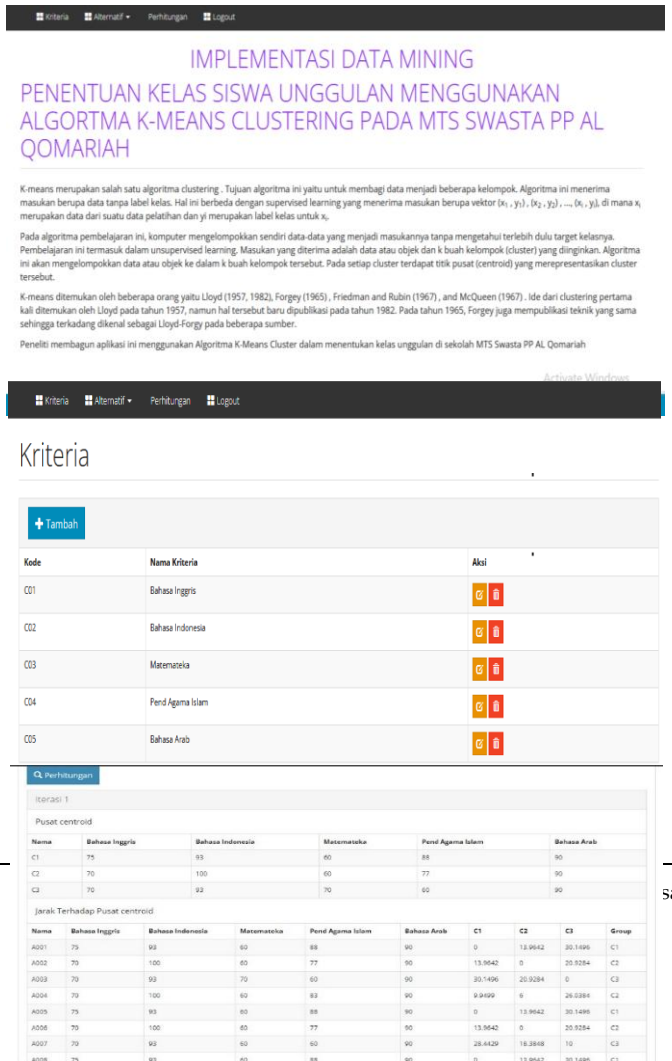
HASIL PENELITIAN

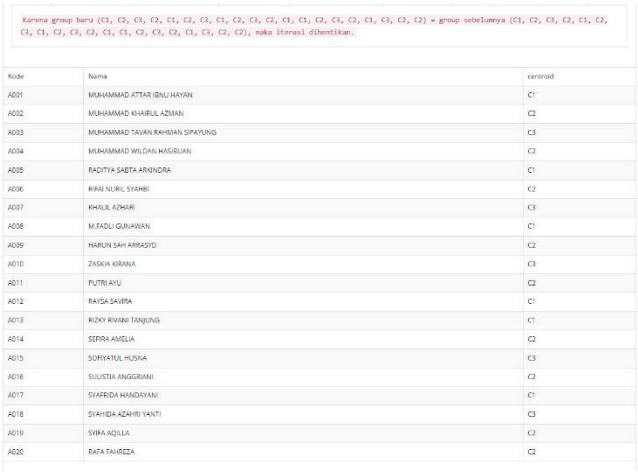
Hasil penelitian berhasil menentukan berapa banyak cluster dan siswa yang dapat dikelompok kedalam cluster C1, cluster C2, dan cluster C3, dengan penjelasan sebagai berikut:

1. Cluster Kelas Siswa Unggul (C1) berdasarkan hasil perhitungan awal atau iterasi pertama terdiri dari 6 data siswa, yaitu nomor 1, 5, 8, 12, 13, 17.
2. Cluster Kelas Siswa Baik (C2) berdasarkan hasil perhitungan awal atau iterasi pertama mencakup 9 data siswa, yaitu 2, 4, 6, 9, 11, 14, 16, 19, 20.
3. Cluster Kelas Siswa Cukup (C3) berdasarkan hasil perhitungan awal atau iterasi pertama mencakup 5 data siswa, yaitu 3, 7, 10, 15, 18.
4. Dari hasil perhitungan iterasi ke-2 tersebut sama dengan hasil perhitungan iterasi ke-1. Oleh karena itu perhitungan dihentikan.

PEMBAHASAN

Bagian ini menyajikan tampilan dari hasil perancangan sistem yang telah dilakukan. Setiap komponen antarmuka yang telah dirancang dijelaskan secara rinci, mencakup halaman home, halaman kriteria, halaman proses, dan halaman hasil analisa.





Karena group baru (C1, C2, C3, C2, C1, C2, C3, C1, C2, C3, C2, C1, C1, C2, C3, C2, C1, C3, C2, C2) = group sebelumnya (C1, C2, C3, C2, C1, C2, C3, C1, C2, C3, C2, C1, C1, C2, C2, C3, C2, C1, C3, C2, C2), maka iterasi dihentikan.

Kode	Nama	centroid
A001	MUHAMMAD ATTAR IBNU HAYAN	C1
A002	MUHAMMAD KHAIRUL AZMAN	C2
A003	MUHAMMAD TAYAN RAHMAN SIPAYUNG	C3
A004	MUHAMMAD WILDAN HASIRUAN	C2
A005	RADITYA SABTA ARKINDRA	C1
A006	RIFAI NURIL SYAHBI	C2
A007	KHALIL AZHARI	C3
A008	M.FADLI GUNAWAN	C1
A009	HARUN SAH ARRASYD	C2
A010	ZASKIA KIRANA	C3
A011	PUTRI AYU	C2
A012	RAYSA SAVIRA	C1
A013	RIZKY RIYANI TANJUNG	C1
A014	SEFIRA AMELIA	C2
A015	SOFIYATUL HUSNA	C3
A016	SULISTIA ANGGRIANI	C2
A017	SYAFRIDA HANDAYANI	C1
A018	SYAHIDA AZAHRI YANTI	C3
A019	SYIFA AQILLA	C2
A020	RAFA FAHREZA	C2

Gambar 2 Antarmuka Hasil Perhitungan

KESIMPULAN

Berdasarkan uraian permasalahan dan pembahasan pada bab sebelumnya, penulis ingin menyampaikan beberapa kesimpulan dan saran yang dapat dijadikan sebagai bahan masukan dan pertimbangan bagi pihak Koperasi. Selama proses perancangan, pembuatan, dan evaluasi, penulis telah menarik beberapa kesimpulan sebagai berikut:

1. Dengan implementasi algoritma k-means clustering dalam penentuan kelas siswa unggulan di sekolah MTS Swasta PP Al Qomariah, sebanyak 20 data siswa yang diuji berhasil dikelompokkan ke dalam 3 cluster, yaitu: kelas siswa unggulan (C1) terdiri dari 6 siswa (data ke-1,5,8,12,13,17), kelas siswa baik (C2) terdiri dari 9 siswa (data ke- 2,4,6,9,11,14,16,19,20), serta juga kelas siswa cukup (C3) terdiri dari 5 siswa (data ke-3,7,10,15,18).
2. Perancangan dan pembangunan sistem menggunakan Algoritma K-Means mampu menghasilkan kelas siswa unggulan di sekolah MTS Swasta PP Al Qomariah ke dalam 3 cluster, sehingga memudahkan pihak sekolah dalam melakukan penentuan kelas siswa unggulan di sekolah MTS Swasta PP Al Qomariah setiap tahunnya.
3. Aplikasi yang dibangun memiliki beberapa pilihan dalam menentukan jumlah pusat cluster dan juga aplikasi bisa menentukan pusat cluster secara mandiri diluar keinginan pengguna sistem.

REFERENCES

Preddy, …, Marpaung, P., Pebrian, I., & Putri, W. (2023). Penerapan Data Mining Untuk Pengelompokan Kepadatan Penduduk Kabupaten Deli Serdang Menggunakan Algoritma K-Means. Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI), 6(2), 64–70.

Hasugian, P. S., & Panjaitan, S. (2018). The Data Mining of Cell Phone Most Interested Using Apriorial Algorithm. INFOKUM, 7(1).

Shah, N., & Shah, K. (2023). Introduction to Data Mining. In Practical Data Mining Techniques and Applications. https://doi.org/10.1201/9781003390220-19 .

P. S. Hasugian and J. Raphita Sagala, "Penerapan Data Mining Untuk Pengelompokan Siswa Berdasarkan Nilai Akademik dengan Algoritma K-Means," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 3, no. 3, 2022.

P. Marpaung and R. F. Siahaan, "Penerapan Algoritma K-Means Clustering Untuk Pemetaan Kepadatan Penduduk Berdasarkan Jumlah Penduduk Kota Medan," *J-SAKTI (Jurnal Sains Komput. dan Inform.*, 2021.