

## **BAB III**

### **ANALISA DAN PERANCANGAN SISTEM**

#### **3.1. Analisa Sistem**

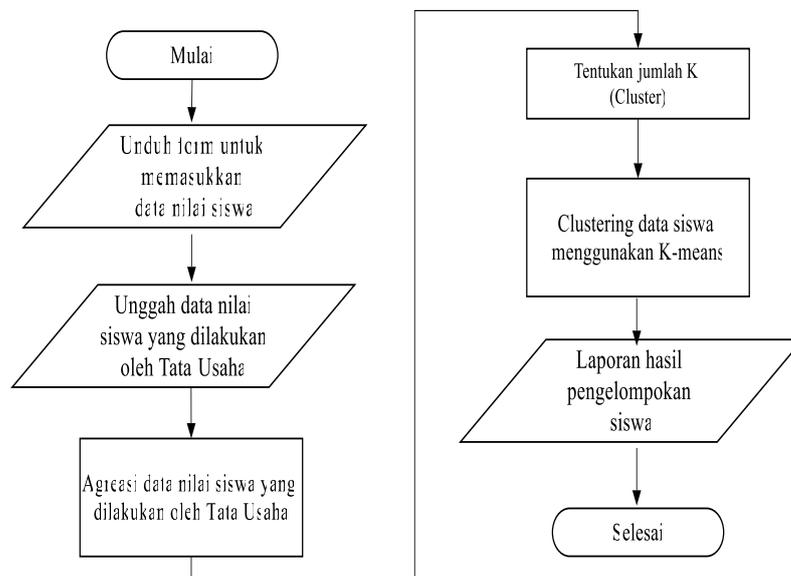
Saat ini dalam pembentukan ruang kelas unggulan masih dilakukan dengan cara mengambil rata-rata tertinggi dari hasil nilai ujian kenaikan kelas, sedangkan sisanya dimasukkan dalam ruang kelas sedang yang dilakukan secara random, menurut pengamatan dan hasil evaluasi bapak/ibu guru yang mengajar di kelas VIII, bahwa ruang kelas H sebagian besar siswanya adalah siswa kurang pintar, sedangkan pada ruang kelas C sebagian besar siswanya adalah siswa yang pintar, hal ini menunjukkan bahwa pembentukan ruang kelas sedang belum berjalan dengan maksimal. Cara mengelompokkan siswa berdasarkan hasil evaluasi belajar siswa yang dikelompokkan per mata pelajaran berikut : Al-Qur'an Hadits, Aqidah, Fiqih, SKI, PKN, B.Indonesia, B.Arab, B.Ingggris, Matematika, IPA, IPS, Seni, Penjas, TIK, B.Jawa, dan PLH. Dengan adanya pengelompokan ini dapat diketahui cluster siswa dengan prestasi akademik tinggi, sedang, dan rendah.

#### **3.2. Hasil Analisis**

Hasil analisis masalah didapatkan bahwa data nilai akademik yang diambil dari hasil evaluasi belajar dari tiap-tiap mata pelajaran dapat diproses menggunakan *Clustering*. Hasilnya berupa informasi yang dapat memudahkan pekerjaan bagi Waka. Kurikulum dalam mengetahui siswa yang mempunyai prestasi akademik baik, sedang dan kurang, yang bisa dijadikan acuan dalam membentuk kondisi kelas yang seimbang. Pengelompokan ini juga diharapkan membantu wali kelas dan guru dalam mengetahui perkembangan anak didiknya dalam bidang akademik, yang nantinya bisa dijadikan acuan oleh guru untuk memperbaiki sistem pengajaran selama ini dan memberikan perlakuan-perlakuan khusus terhadap tingkatan-tingkatan kemampuan siswa, sehingga proses pembelajaran dapat tercapai dalam seluruh kelas dan dalam semua tingkatan kemampuan siswa.

### 3.2.1. Deskripsi Sistem

Sistem yang dibangun adalah pengelompokan siswa MTs Negeri Gresik berdasarkan nilai akademik. Sistem yang dibangun ini bertujuan mengelompokkan siswa berdasarkan nilai akademik dari hasil evaluasi belajar siswa. Adapun diagram alirnya dapat digambarkan pada Gambar 3.1 :



**Gambar 3.1** *Flowchart System*

Keterangan proses yang terjadi pada gambar 3.1:

1. Tata Usaha (TU) mengunduh *form* isian data nilai siswa dari halaman *Download Form* dan memasukkan data nilai siswa berdasarkan data dari wali kelas/guru mata pelajaran.
2. TU memasukkan atau mengunggah *form* data nilai siswa kedalam sistem.
3. TU melakukan proses agregasi data nilai siswa. Maksud agregasi ini digunakan untuk meringkas data awal nilai siswa, dari beberapa data tersebut nantinya akan diagregasikan sehingga didapatkan satu data akhir.
4. Menentukan jumlah *cluster* (K).

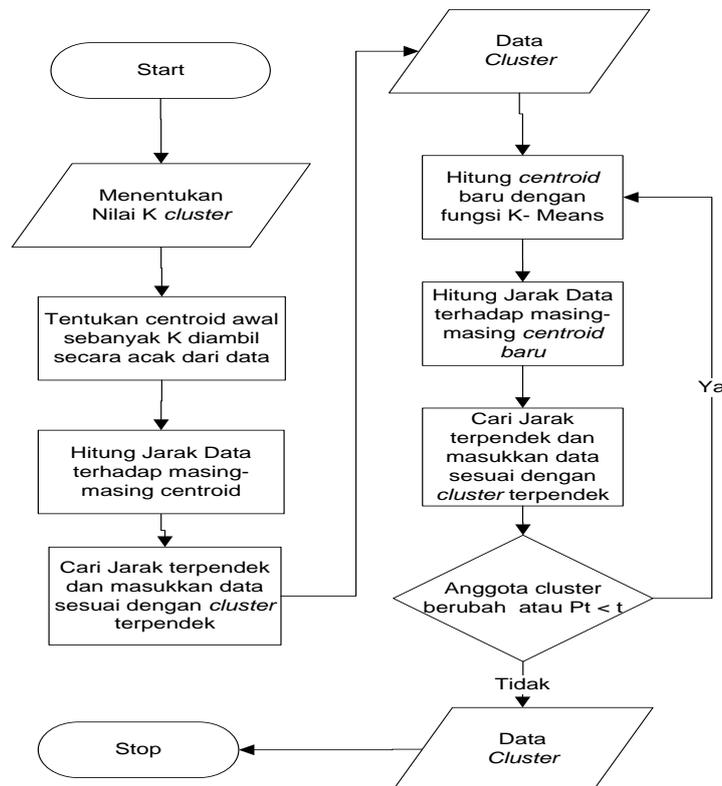
5. Sistem akan memulai melakukan proses pengelompokan menggunakan metode *K-means*.
6. Sistem menampilkan hasil akhir dari *Clustering*, berupa laporan nama-nama siswa yang masuk ke dalam kelompok prestasi akademik tinggi, sedang, dan rendah.

### 3.2.2. Kebutuhan Data

Kebutuhan data yang dijadikan inputan dalam sistem pengelompokan siswa ini adalah data nilai siswa kelas VIII MTs Negeri Gresik tahun pelajaran 2016-2017 sebanyak 314 siswa, yang meliputi nilai nilai ulangan harian, nilai penerapan, nilai ujian tengah semester, dan nilai semester tiap-tiap mata pelajaran berikut : Al-Qur'an Hadits, Aqidah, Fiqih, SKI, PKN, B.Indonesia, B.Arab, B.Ingggris, Matematika, IPA, IPS, Seni, Penjas, TIK, B.Jawa, dan PLH.

### 3.2.3. Proses Perhitungan *K-means*

*K-means* merupakan salah satu metode pengelompokan data nonhierarki yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan varisai dalam suatu kelompok dan memaksimalkan variasi antar kelompok. *K-means* merupakan salah satu metode pengelompokan data nonhierarki (Prasetyo,2012). Berikut ini adalah alur diagram dari algoritma *k-means* digambarkan pada Gambar 3.2 :



**Gambar 3.2** Diagram alir Algoritma *K-means*

Pada Gambar 3.2 menjelaskan algoritma *K-means* yang digunakan untuk pengelompokan prestasi akademik siswa. Pertama menentukan nilai *K cluster* kemudian menentukan nilai *centroid* awal secara acak, selanjutnya menentukan jarak terdekat terhadap *centroid* pada setiap data dalam kasus ini digunakan rumus *Euclidian distance*. Kemudian cari jarak terdekat data dan masukkan data ke *cluster* sesuai dengan *centroid* terdekat. Kemudian mengecek anggota *cluster* apakah anggota *cluster* berubah atau apabila ada perubahan nilai fungsi objektif diatas nilai *threshol*d. Apabila ada perubahan mengulangi langkah menghitung *centroid* baru dengan fungsi *K-means* hingga anggota suatu cluster tidak berubah.

### 3.3. Representasi Data

Data awal yang digunakan dalam pengelompokan siswa berdasarkan nilai akademik siswa MTs Negeri Gresik dari ruang C sebanyak 20 siswa, terdapat pada lampiran A (Data Siswa).

#### 3.3.1 Agregasi data.

Dari proses agregasi ini, diperoleh bobot nilai akhir, sebagai berikut :

Nilai Akhir =

$$(60\% \times (\text{jumlah nilai harian})) + (40\% \times (\text{jumlah nilai ujian}))$$

Contoh perhitungan agregasi data pada mata pelajaran Al-Qur'an Hadits.

$$\begin{aligned} \text{Nilai Akhir} &= (60\% \times (80 + 75 + 80 + 80)) + (40\% \times (70 + 75)) \\ &= 76 \end{aligned}$$

Agregasi data diperoleh pada data awal siswa dari ruang C sebanyak 20 siswa, terdapat pada lampiran B.

#### 3.3.2 K-means

Dalam tahap ini akan dilakukan proses utama yaitu segmentasi data nilai yang diakses dari database yaitu dengan metode *K-means*. Berikut ini merupakan asumsi bahwa inputan adalah jumlah data set sebanyak  $N = 20$  data yaitu jumlah banyaknya siswa dan jumlah inisialisasi centroid  $K = 3$

Selanjutnya akan digunakan algoritma pengelompokan *K-means* untuk mengelompokan data yang ada. Data yang ada akan di kelompokkan menjadi beberapa kelompok, adapun langkah dari pengelompokan data adalah:

1. Penentuan pusat awal *cluster*, untuk penentuan awal *cluster* dilakukan secara acak (random) dari data nilai siswa :

**Tabel 3.1** Pusat *cluster* awal

Data ke 5	C1	80	88	79	87	80	79	81	70	76	79	68	70	73	83	80	73
Data ke 11	C2	76	79	81	87	79	75	76	69	76	77	74	64	74	83	83	77
Data ke 18	C3	84	85	85	88	85	85	86	81	87	83	87	64	80	84	81	77

2. Menghitung jarak setiap data yang ada terhadap setiap pusat *cluster* menggunakan rumus *Euclidian Distance*, seperti pada (2.1).

Jarak nilai terhadap C1 :

$$D1c1 = \sqrt{\begin{matrix} (76-80)^2 + (78-88)^2 + (76-79)^2 + (81-87)^2 + (78-80)^2 + (74-79)^2 + (70-81)^2 + \\ (66-70)^2 + (75-76)^2 + (78-79)^2 + (72-68)^2 + (73-70)^2 + (74-73)^2 + (85-83)^2 + \\ (83-80)^2 + (73-73)^2 \end{matrix}}$$

$$=19,04$$

Jarak nilai terhadap C2 :

$$D1c2 = \sqrt{\begin{matrix} (76-76)^2 + (78-79)^2 + (76-81)^2 + (81-87)^2 + (78-79)^2 + (74-75)^2 + (70-76)^2 + \\ (66-69)^2 + (75-76)^2 + (78-77)^2 + (72-74)^2 + (73-64)^2 + (74-74)^2 + (85-83)^2 + \\ (83-83)^2 + (73-77)^2 \end{matrix}}$$

$$=14,99$$

Jarak nilai terhadap C3 :

$$D1c3 = \sqrt{\begin{matrix} (76-84)^2 + (78-85)^2 + (76-85)^2 + (81-88)^2 + (78-85)^2 + (74-85)^2 + (70-86)^2 + \\ (66-81)^2 + (75-87)^2 + (78-83)^2 + (72-87)^2 + (73-64)^2 + (74-80)^2 + (85-84)^2 + \\ (83-81)^2 + (73-77)^2 \end{matrix}}$$

$$=37,61$$

Perhitungan jarak data terhadap *cluster* selanjutnya sama seperti perhitungan jarak data terhadap *cluster* sebelumnya.

Maka hasil yang diperoleh yaitu terlihat pada tabel 3.2 sebagai berikut ini :

**Tabel 3.2** Tabel jarak dan *Cluster* pada iterasi 1

No	$d_nC1$	$d_nC2$	$d_nC3$	Jarak terdekat	cluster
1	19.04	14.99	37.61	14.99	2
2	12.69	13.26	34.72	12.69	1
3	15.83	13.35	35.46	13.35	2
4	11.17	19.03	24.40	11.17	1
5	0.00	15.66	29.40	0.00	1
6	17.11	14.67	36.59	14.67	2
7	8.96	13.45	27.35	8.96	1
8	12.44	15.60	25.13	12.44	1
9	19.54	17.48	21.04	17.48	2
10	14.36	7.07	27.63	7.07	2
11	15.66	0.00	28.92	0.00	2
12	19.08	15.65	20.66	15.65	2
13	15.96	10.44	24.78	10.44	2
14	12.84	13.48	25.15	12.84	1
15	15.04	13.40	33.58	13.40	2
16	13.52	14.56	24.90	13.52	1
17	24.01	26.61	15.98	15.98	3

No	$d_nC1$	$d_nC2$	$d_nC3$	Jarak terdekat	cluster
18	29.40	28.92	0.00	0.00	3
19	19.23	19.72	18.39	18.39	3
20	16.80	19.98	19.82	16.80	1

Maka hasil dari tabel 3.2 diatas, untuk *cluster* pertama ada 8 anggota siswa, *cluster* ke dua mempunyai 9 anggota siswa dan *cluster* ketiga mempunyai 3 anggota siswa.

### Iterasi 2

Setelah mengetahui anggota tiap-tiap *cluster* kemudian kita menentukan *cluster* baru yang dihitung berdasarkan data anggota tiap *cluster* yang sesuai dengan rumus pusat anggota *cluster* baru sehingga didapatkan perhitungan.

Yang terlihat pada tabel 3.3 Sebagai berikut :

**Tabel 3.3** pusat *cluster* baru

C1	79.25	85.81	79.54	86.74	79.58	78.58	79.86	71.34	75.93	80.16	74.11	71.51	75.06	84.33	81.79	75.33
C2	76.96	79.79	78.46	85.23	78.77	75.32	76.38	70.84	75.37	78.36	75.85	69.99	73.73	83.33	82.15	74.58
C3	83.50	86.72	83.43	88.23	85.05	83.18	83.03	76.62	81.35	82.90	84.25	68.50	79.67	83.22	84.15	76.28

Setelah menentukan titik pusat yang baru kita akan menghitung jarak di setiap data terhadap pusat *cluster*. Kemudian menentukan *cluster* dengan jarak terdekat pada masing-masing data. Berikut ini tabel 3.4 yaitu sebagai berikut :

**Tabel 3.4** Tabel jarak dan *Cluster* pada iterasi 2

No	$d_nC1$	$d_nC2$	$d_nC3$	Jarak terdekat	cluster
1	17.07	11.18	30.67	11.18	2
2	11.55	8.97	27.37	8.97	2
3	14.09	9.22	28.21	9.22	2
4	8.67	15.07	18.51	8.67	1
5	7.88	13.58	22.54	7.88	1
6	16.11	10.66	30.45	10.66	2
7	6.94	9.27	20.78	6.94	1
8	7.81	11.86	17.40	7.81	1
9	13.94	14.60	15.23	13.94	1
10	10.34	6.50	20.98	6.50	2
11	13.02	8.04	23.44	8.04	2
12	13.45	13.32	16.32	13.32	2
13	10.03	6.26	18.86	6.26	2

No	$d_nC1$	$d_nC2$	$d_nC3$	Jarak terdekat	cluster
14	8.01	11.56	18.36	8.01	1
15	12.29	8.07	26.29	8.07	2
16	8.17	12.95	16.86	8.17	1
17	18.70	24.37	8.82	8.82	3
18	25.09	28.47	10.13	10.13	3
19	14.69	17.47	10.26	10.26	3
20	11.96	16.76	11.05	11.05	3

Maka hasil dari tabel 3.4 diatas, untuk *cluster* pertama ada 7 anggota siswa, *cluster* ke dua mempunyai 9 anggota siswa dan *cluster* ketiga mempunyai 4 anggota siswa. Setelah mengetahui hasil *cluster* masih mengalami perubahan maka kita mengulangi langkah tersebut sampai posisi *cluster* tidak mengalami perubahan.

### Iterasi 3

Setelah mengetahui anggota tiap-tiap *cluster* kemudian kita menentukan *cluster* baru yang dihitung berdasarkan data anggota tiap *cluster* yang sesuai dengan rumus pusat anggota *cluster* baru sehingga didapatkan perhitungan.

Yang terlihat pada tabel 3.5 Sebagai berikut :

**Tabel 3.5** pusat *cluster* baru

C1	79.18	85.50	79.41	87.59	78.77	78.33	81.76	71.36	76.43	80.32	75.51	70.93	74.70	84.39	81.86	75.21
C2	76.63	79.92	78.26	84.59	78.56	75.18	75.05	70.34	74.98	78.04	74.30	70.32	73.72	83.54	81.95	74.57
C3	83.31	86.73	83.14	87.83	85.56	82.79	81.89	76.38	80.01	82.63	82.76	69.54	79.18	82.93	83.89	76.26

Setelah menentukan titik pusat yang baru kita akan menghitung jarak di setiap data terhadap pusat *cluster*. Kemudian menentukan *cluster* dengan jarak terdekat pada masing-masing data. Berikut ini tabel 3.6 yaitu sebagai berikut :

**Tabel 3.6** Tabel jarak dan *Cluster* pada iterasi 3

No	$d_nC1$	$d_nC2$	$d_nC3$	Jarak terdekat	cluster
1	18.53	9.32	28.82	9.32	2
2	13.26	7.16	25.37	7.16	2
3	15.69	7.49	26.18	7.49	2
4	7.95	15.93	17.15	7.95	1
5	9.05	13.42	20.71	9.05	1
6	17.19	9.25	28.60	9.25	2
7	7.76	9.58	18.93	7.76	1

No	$d_nC1$	$d_nC2$	$d_nC3$	Jarak terdekat	cluster
8	8.18	12.53	15.89	8.18	1
9	11.77	16.82	14.85	11.77	1
10	10.69	7.23	19.45	7.23	2
11	13.22	8.39	22.11	8.39	2
12	12.02	15.24	15.57	12.02	1
13	9.89	7.94	17.32	7.94	2
14	6.71	12.77	17.12	6.71	1
15	13.73	6.76	24.32	6.76	2
16	7.80	13.96	15.36	7.80	1
17	17.69	26.00	9.08	9.08	3
18	23.96	30.26	11.83	11.83	3
19	14.92	18.75	8.60	8.60	3
20	12.76	17.79	8.29	8.29	3

Maka hasil dari tabel 3.6 diatas, untuk *cluster* pertama ada 8 anggota siswa, *cluster* ke dua mempunyai 8 anggota siswa dan *cluster* ketiga mempunyai 4 anggota siswa. Setelah mengetahui hasil *cluster* masih mengalami perubahan maka kita mengulangi langkah tersebut sampai posisi *cluster* tidak mengalami perubahan.

#### Iterasi 4

Setelah mengetahui anggota tiap-tiap *cluster* kemudian kita menentukan *cluster* baru yang dihitung berdasarkan data anggota tiap *cluster* yang sesuai dengan rumus pusat anggota *cluster* baru sehingga didapatkan perhitungan. Yang terlihat pada tabel 3.7 Sebagai berikut :

**Tabel 3.7** pusat *cluster* baru

C1	78.69	84.82	79.51	87.64	78.83	78.54	81.93	72.24	76.46	80.26	76.23	70.65	74.43	84.38	81.90	75.30
C2	76.80	79.91	78.02	84.16	78.47	74.58	74.04	69.34	74.77	77.82	73.43	70.52	73.88	83.45	81.92	74.41
C3	83.31	86.73	83.14	87.83	85.56	82.79	81.89	76.38	80.01	82.63	82.76	69.54	79.18	82.93	83.89	76.26

Setelah menentukan titik pusat yang baru kita akan menghitung jarak di setiap data terhadap pusat *cluster*. Kemudian menentukan *cluster* dengan jarak terdekat pada masing-masing data. Berikut ini tabel 3.8 yaitu sebagai berikut :

**Tabel 3.8** Tabel jarak dan *Cluster* pada iterasi 4

No	d <sub>n</sub> C1	d <sub>n</sub> C2	d <sub>n</sub> C3	Jarak terdekat	cluster
1	18.82	7.81	28.82	7.81	2
2	13.74	6.30	74.39	6.30	2
3	16.07	6.09	26.18	6.09	2
4	8.29	17.00	17.15	8.29	1
5	10.07	13.77	20.71	10.07	1
6	17.37	8.36	28.60	8.36	2
7	8.09	10.36	18.93	8.09	1
8	8.31	13.52	15.89	8.31	1
9	10.82	18.42	14.85	10.82	1
10	10.46	8.14	19.45	8.14	2
11	12.95	9.02	22.11	9.02	2
12	10.52	17.14	15.57	10.52	1
13	8.98	9.63	17.32	8.98	1
14	6.68	13.97	17.12	6.68	1
15	14.11	5.82	24.32	5.82	2
16	8.12	14.95	15.36	8.12	1
17	17.36	27.38	9.08	9.08	3
18	23.23	31.77	12.34	12.34	3
19	14.77	19.72	8.60	8.60	3
20	12.63	18.87	8.29	8.29	3

Maka hasil dari tabel 3.8 diatas, untuk *cluster* pertama ada 9 anggota siswa, *cluster* ke dua mempunyai 7 anggota siswa dan *cluster* ketiga mempunyai 4 anggota siswa. Setelah mengetahui hasil *cluster* masih mengalami perubahan maka kita mengulangi langkah tersebut sampai posisi *cluster* tidak mengalami perubahan.

### Iterasi 5

Setelah mengetahui anggota tiap-tiap *cluster* kemudian kita menentukan *cluster* baru yang dihitung berdasarkan data anggota tiap *cluster* yang sesuai dengan rumus pusat anggota *cluster* baru sehingga didapatkan perhitungan.

Yang terlihat pada tabel 3.9 Sebagai berikut :

**Tabel 3.9** pusat *cluster* baru

C1	78.53	84.31	79.66	87.41	78.84	78.34	81.42	72.46	76.29	80.06	76.61	70.53	74.28	84.29	81.72	75.31
C2	76.74	79.86	77.61	83.96	78.40	74.26	73.58	68.65	74.74	77.74	72.54	70.65	73.99	83.42	82.16	74.26

C3	83.31	86.73	83.14	87.83	85.56	82.79	81.89	76.38	80.01	82.63	82.76	69.54	79.18	82.93	83.89	76.26
----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Setelah menentukan titik pusat yang baru kita akan menghitung jarak di setiap data terhadap pusat *cluster*. Kemudian menentukan *cluster* dengan jarak terdekat pada masing-masing data. Berikut ini tabel 3.10 yaitu sebagai berikut:

**Tabel 3.10** Tabel jarak dan *Cluster* pada iterasi 5

No	$d_nC1$	$d_nC2$	$d_nC3$	Jarak terdekat	cluster
1	18.33	6.87	28.82	6.87	2
2	13.47	5.70	25.37	5.70	2
3	15.64	5.26	26.18	5.26	2
4	8.80	17.72	17.15	8.80	1
5	10.51	13.98	20.71	10.51	1
6	16.84	8.06	28.60	8.06	2
7	7.93	11.03	18.93	7.93	1
8	8.48	14.15	15.89	8.48	1
9	10.69	19.48	14.85	10.69	1
10	9.93	8.88	19.45	8.88	2
11	12.38	9.62	22.11	9.62	2
12	10.00	18.40	15.57	10.00	1
13	7.98	11.01	17.32	7.98	1
14	6.80	14.81	17.12	6.80	1
15	13.64	5.68	24.32	5.68	2
16	8.23	15.77	15.36	8.23	1
17	17.66	28.37	9.08	9.08	3
18	23.24	32.88	12.34	12.34	3
19	14.69	20.53	8.60	8.60	3
20	12.61	19.75	8.29	8.29	3

Maka hasil dari tabel 3.10 diatas, untuk *cluster* pertama ada 9 anggota siswa, *cluster* ke dua mempunyai 7 anggota siswa dan *cluster* ketiga mempunyai 4 anggota siswa. Setelah mengetahui hasil *cluster* masih mengalami perubahan maka kita mengulangi langkah tersebut sampai posisi *cluster* tidak mengalami perubahan.

Karena ke 5 posisi *cluster* tidak berubah maka iterasi dihentikan dan hasil akhir yang diperoleh adalah 3 *cluster* yaitu :

- *cluster* pertama ada 9 anggota siswa, yaitu data ke : 4,5,7,8,9,12,13,14,16.

- *cluster* ke dua mempunyai 7 anggota siswa, yaitu data ke : 1,2,3,6,10,11,15.
- *cluster* ketiga mempunyai 4 anggota siswa, yaitu data ke : 17,18,19,20.

Dalam tiga *cluster* tersebut bisa didapatkan *cluster* nilai terkecil, *cluster* nilai sedang, dan *cluster* nilai terbesar yaitu dengan melakukan rata-rata semua fitur dari setiap nilai *centroid* pada iterasi terakhir. Berikut tabel nilai rata-rata *cluster* terakhir :

**Tabel 3.11** Hasil rata-rata *centroid* terakhir

C1	78,75
C2	76,41
C3	81,55

Pada tabel 3.11 menjelaskan bahwa hasil rata-rata setiap *centroid*, yang diambil dari data *centroid* akhir dan yang mendapatkan nilai rata-rata terbesar adalah pada *centroid* 3 yaitu 81,55 dan yang mendapatkan nilai rata-rata sedang adalah pada *centroid* 2 yaitu 76,41 dan yang mendapatkan nilai rata-rata terkecil adalah pada *centroid* 1 yaitu 78,75. Maka didapatkan kesimpulan bahwa siswa yang mengikuti *cluster* 1 merupakan siswa dengan prestasi akademik kurang dan *cluster* 2 merupakan siswa dengan prestasi akademik cukup dan *cluster* 3 merupakan siswa dengan prestasi akademik baik. Proses perangkingan *cluster* ini hanya bisa dilakukan jika  $K=3$ , jika  $K$  tidak sama dengan 3 maka tidak bisa dilakukan proses perangkingan tersebut dan hanya bisa dilakukan proses pengelompokan biasa

### 3.3.3 Evaluasi *Cluster Davies Bouldin Index*

Evaluasi *cluster* yang akan digunakan dalam sistem ini adalah evaluasi validitas internal, yakni dalam evaluasi hasil *cluster* tanpa menggunakan informasi dari luar/eksternal. Evaluasi validitas *cluster* akan membandingkan hasil *cluster* dengan nilai  $K=3$  dengan metode validitas Davies Bouldin Index yang diperkenalkan oleh David L. Davies dan Donald W. Bouldin pada tahun

1972. Dimana nilai DBI yang terkecil maka *cluster* tersebut yang paling bagus / valid.

Perhitungan SSB dari *centroid* terakhir yang didapatkan dalam proses akhir *Clustering* seperti pada hasil tabel 3.11 diatas adalah sebagai berikut :

- Jarak antara *centroid* 1 dengan *centroid* 2, seperti pada (2.3).

$$SSB_{1,2} = \sqrt{(78,53 - 76,74)^2 + (84,31 - 79,86)^2 + (79,66 - 77,61)^2 + (87,41 - 83,96)^2 + (78,84 - 78,4)^2 + (78,34 - 74,26)^2 + (81,42 - 73,58)^2 + (72,46 - 68,65)^2 + (76,29 - 74,74)^2 + (80,06 - 77,74)^2 + (76,61 - 72,54)^2 + (70,53 - 70,65)^2 + (74,28 - 73,99)^2 + (84,29 - 83,42)^2 + (81,72 - 82,16)^2 + (75,31 - 74,26)^2}$$

$$= 13$$

- Jarak antara *centroid* 1 dengan *centroid* 3

$$SSB_{1,3} = 15$$

- Jarak antara *centroid* 2 dengan *centroid* 3

$$SSB_{2,3} = 24$$

Perhitungan SSW dari hasil *cluster* terakhir yang didapatkan dalam proses akhir *Clustering* seperti pada hasil tabel 3.11 diatas adalah sebagai berikut :

- Rata-rata jarak antara data *cluster* 1 dengan *centroid* 1, seperti pada (2.4).

$$SSW_1 = 8,82$$

- Rata-rata jarak antara data *cluster* 2 dengan *centroid* 2

$$SSW_2 = 7,15$$

- Rata-rata jarak antara data *cluster* 3 dengan *centroid* 3

$$SSW_3 = 9,58$$

Perhitungan R didapatkan dengan menghitung SSB dan SSW seperti dibawah ini :

- $R_{12} = \frac{SSW_1 + SSW_2}{SSB_{1,2}} = 1,27$

- $R_{13} = \frac{SSW_1 + SSW_3}{SSB_{1,3}} = 1,26$

- $R_{23} = \frac{SSW_2 + SSW_3}{SSB_{2,3}} = 0,69$

Untuk Mencari DBI maka hasil perhitungan R dipilih yang tertinggi kemudian dirata-rata seperti berikut :

**Tabel 3.12** Tabel Perhitungan DBI

R	Data ke- i			R Max
	1	2	3	
1	0	1,27	1,26	1,27
2	1,27	0	0,69	1,27
3	1,26	0,69	0	1,26

$$DBI = \frac{1}{3} (1,27 + 1,27 + 1,26) = 1,27$$

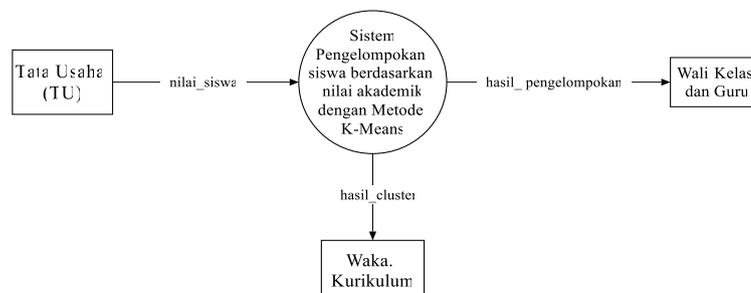
Maka dari hasil *cluster* diatas nilai validitas *cluster* dengan DBI adalah 1,27. Untuk pemilihan hasil *cluster* yang terbaik adalah dengan memilih *cluster* yang memiliki nilai DBI terkecil dalam studi kasus ini percobaan *cluster* dilakukan sebanyak 10 kali dengan nilai K=3.

### 3.4. Perancangan Sistem

Perancangan sistem merupakan tahapan setelah melakukan analisis dari pengembangan sistem, pendefinisian dari kebutuhan-kebutuhan fungsional dan persiapan untuk rancang bangun implementasi dan gambaran bagaimana suatu sistem dapat terbentuk.

#### 3.4.1. Diagram Konteks

Diagram konteks merupakan diagram yang menjelaskan secara keseluruhan proses utama dalam sebuah sistem. Diagram tersebut menjelaskan apa yang dimasukkan dan yang diterima oleh pengguna sistem.

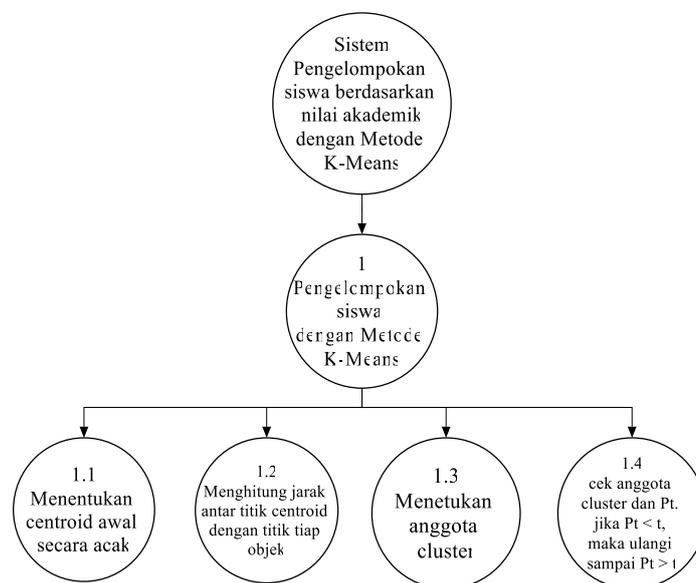


**Gambar 3.3** Diagram konteks

Keterangan gambar 3.3 :

1. Tata Usaha berperan dalam memasukkan nilai prestasi akademik siswa. dan mendapatkan laporan hasil *cluster* secara details.
2. Waka kurikulum mendapatkan laporan hasil *cluster* secara details.
3. Guru dan Wali Kelas yaitu pihak yang hanya dapat melihat hasil laporan pengelompokan siswa yang telah diproses di aplikasi tersebut.

### 3.4.2. Diagram berjenjang



**Gambar 3.4** Diagram Berjenjang

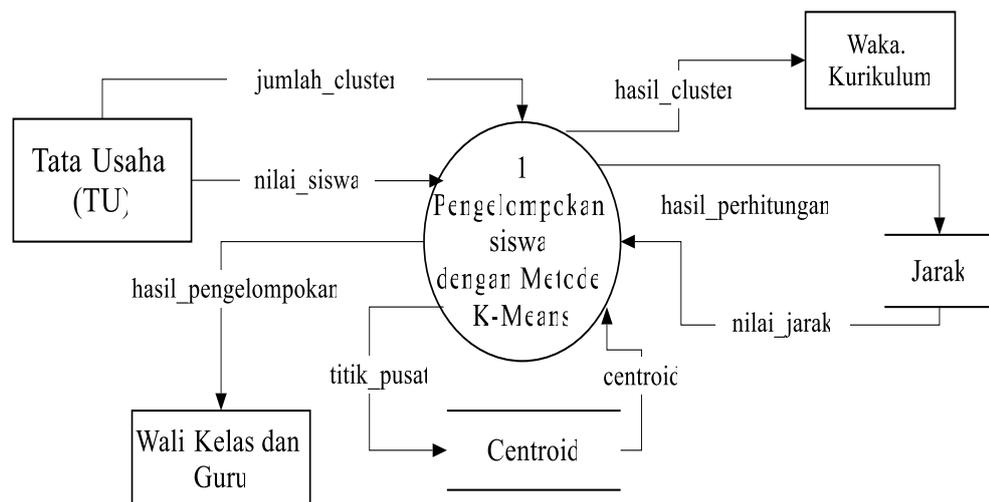
Keterangan:

- *Top Level* : Pengeolompokan Siswa Berdasarkan Nilai Akademik Menggunakan metode *K-means*
- *Level 0* : Merupakan hasil *break down* dari proses keseluruhan dari sistem Pengeolompokan Siswa Berdasarkan Nilai Akademik Menggunakan metode *K-means* yang dibagi menjadi beberapa sub proses antara lain;
  1. Pengelompokan siswa Berdasarkan Nilai Akademik dengan *K-means*.

- Level 1 : Merupakan sub proses dari Level 0 sistem Pengeolompokan Siswa Berdasarkan Nilai Akademik Menggunakan metode *K-means* yang sudah dibagi menjadi beberapa sub proses antara lain;
  1. Pengelompokan siswa dengan *K-means* yang meliputi :
    - 1.1 Menentukan *centroid* awal secara acak.
    - 1.2 Menghitung jarak antar titik *centroid* dengan titik tiap *cluster*
    - 1.3 Menentukan anggota *cluster*.
    - 1.4 Cek anggota *cluster* dan  $Pt < t$ , Maka ulangi sampai  $Pt > t$

### 3.4.3. Data Flow Diagram

Berikut gambar DFD Level 0 :

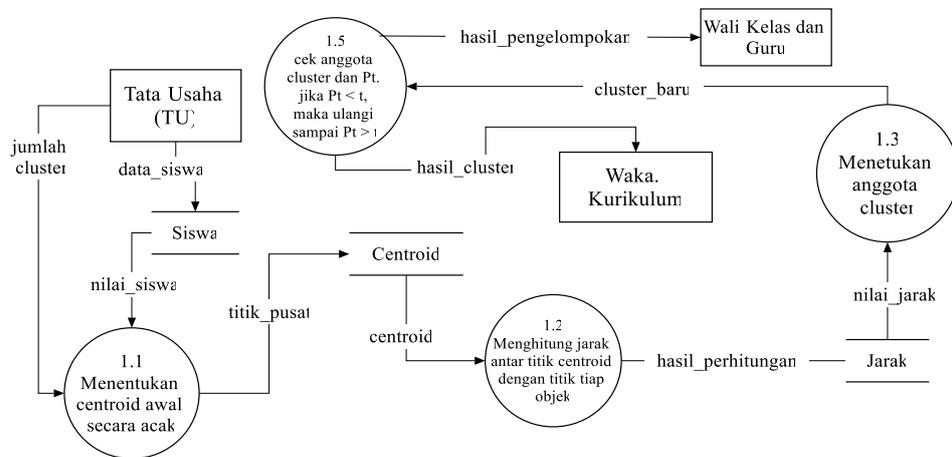


**Gambar 3.5** DFD Level 0

Keterangan dari gambar 3.5 adalah sebagai berikut :

1. Proses 1 adalah perhitungan pengelompokan siswa dengan *K-means* yaitu melakukan proses perhitungan *Clustering*. Waka. Kurikulum mendapatkan hasil *cluster* yang didapatkan dari pengelompokan nilai baik, sedang, dan kurang. Guru dan Wali kelas, pihak yang dapat melihat hasil pengelompokan ruangan.

Berikut Gambar DFD level 1:



**Gambar 3.6** DFD Level 1

Keterangan dari gambar 3.6 :

- 1.1 Tata Usaha memasukkan nilai  $k$  yaitu jumlah *cluster* yang diinginkan sesuai instruksi dari Waka. Kurikulum, kemudian sistem akan memilih *centroid* secara acak dari data nilai pada tabel Siswa sebanyak  $k$  dan akan disimpan di tabel *centroid*.
- 1.2 Kemudian dilanjutkan dengan proses menghitung jarak tiap titik *centroid* dengan titik tiap objek yang didapat dari data nilai pada tabel Siswa.
- 1.3 Setelah menghitung jarak, nilai jarak akan dikelompokkan sesuai jarak terpendek untuk menentukan anggota *cluster*.
- 1.4 Proses selanjutnya adalah Memeriksa anggota *cluster* dan perubahan fungsi objektif, jika posisi *cluster* berubah atau perubahan fungsi objektif  $> T$  maka proses diulangi hingga posisi *cluster* tidak berubah atau perubahan fungsi objektif  $< T$ , kemudian dilaporkan kepada Tata Usaha Waka. Kurikulum, guru dan Walikelas. Didapatkan pihak dari Waka. Kurikulum hasil cluster pengelompokan nilai baik, sedang, dan kurang. Pihak Wali Kelas dan Guru mendapat hasil pengelompokan ruangan.

### 3.5 Kebutuhan Perancangan sistem

Dalam proses pembuatan sistem ini membutuhkan beberapa komponen yang dibagi menjadi kebutuhan perangkat lunak (*software*) dan kebutuhan perangkat keras (*hardware*).

#### 3.5.1. Spesifikasi Perangkat Lunak

Adapun perangkat lunak yang dibutuhkan dalam pembangunan aplikasi tersebut adalah sebagai berikut :

1. Aplikasi browser Firefox Mozilla atau Google Chrome yang berfungsi untuk mengakses sistem pengelompokan siswa dengan metode *K-means* yang berbasis *web*.
2. Bahasa Pemrograman HTML untuk membuat aplikasi berbasis *web*.
3. MySQL adalah sebuah perangkat lunak sistem manajemen basis data SQL yang berfungsi sebagai pengolahan data atau *database*.
4. Apache yang berfungsi sebagai *web server*.
5. XAMPP adalah aplikasi yang berisi semua aplikasi *server* yang berfungsi untuk membuat *server* lokal atau *localhost*.

#### 3.5.2. Spesifikasi Perangkat Keras

Perangkat keras adalah komponen fisik peralatan yang membentuk sistem komputer, serta peralatan lain yang mendukung komputer dalam menjalankan tugasnya. Adapun perangkat keras yang dibutuhkan untuk menjalankan aplikasi ini yaitu :

1. Prosesor intel Pentium dual core
2. RAM 1GB
3. Monitor
4. Keyboard
5. Mouse

### 3.6 Perancangan Basis Data

Dalam proses ini akan menjelaskan tentang kebutuhan *database* atau tabel-tabel yang digunakan untuk keperluan sistem yang akan dibangun.

#### 3.6.1. Table User

Tabel *user* ini digunakan untuk menyimpan *id\_user*, *nama*, *username*, *password*, dan *status* dengan *id\_user* sebagai *primary key* yang menunjukkan urutan dari semua *user* yang terdapat di *database*. Struktur tabel *user* dapat dilihat pada tabel 3.13 :

**Tabel 3.13** Tabel *user*

No	Nama	Tipe	Panjang	Keterangan
1	<i>Id_user</i>	Int	5	primary key
2	Nama	Varchar	25	
3	<i>Username</i>	Varchar	25	
4	Password	Varchar	25	
5	Status	Int		

Variabel pada tabel *user* meliputi :

- *Field id\_user* merupakan nama *id\_user*.
- *Field nama* merupakan nama *user*.
- *Field username* merupakan nama *login*.
- *Field password* merupakan *password login*.
- *Field status* merupakan kode *user*

Status merupakan kode *user*, dalam tabel *user* ini hak akses dibedakan menjadi dua yakni *admin* dan *user*. Untuk *admin* kolom level berisi 0, sedangkan untuk *user* kolom level berisi 1.

#### 3.6.2. Table Siswa

Tabel siswa ini digunakan untuk menyimpan No Induk, Nama, ruang, kode mapel, NH1-4, UTS dan UAS, dengan *No\_induk* sebagai *primary key*. Struktur dari tabel ini dapat dilihat pada tabel 3.14 :

**Tabel 3.14** Tabel Siswa

No	Nama	Tipe	Panjang	Keterangan
1	No_Induk	Int	5	primary key
2	Nama	Varchar	25	
3	Ruang	Varchar	5	
4	Kode Mapel	Int	10	
5	NH1	Double	10	
6	NH2	Double	10	
7	NH3	Double	10	
8	NH4	Double	10	
9	UTS	Double	10	
10	UAS	Double	10	

Tabel siswa berfungsi untuk memberikan informasi data nilai siswa.

Variabel pada tabel meliputi :

- *Field* No Induk merupakan nama No induk setiap siswa.
- *Field* Nama merupakan nama siswa.
- *Field* Ruang merupakan ruang kelas dari siswa
- *Field* Kode\_Mapel merupakan kode dari Mata Pelajaran.
- *Field* NH1-4, UTS, dan UAS merupakan nilai dari tiap-tiap mata pelajaran.

### 3.6.3. Table *Centroid* Awal

Tabel *Centroid* awal ini digunakan untuk menyimpan *id\_centroid\_awal*, *No\_Induk*, *Kode\_Mapel*, *C\_NH1-4*, *C\_UTS*, dan *C\_UAS* dengan *id\_centroid\_awal* sebagai *primary key*. Struktur dari tabel ini dapat dilihat pada tabel 3.15 :

**Tabel 3.15** Tabel Centroid Awal

No	Nama	Tipe	Panjang	Keterangan
1	Id_centroidA	Int	5	primary key
2	No_Induk	Int	10	
3	Kode_Mapel	Int	10	
4	C_NH1	Double	10	
5	C_NH2	Double	10	
6	C_NH3	Double	10	

No	Nama	Tipe	Panjang	Keterangan
7	C_NH4	Double	10	
8	C_UTS	Double	10	
9	C_UAS	Double	10	

Tabel *centroid* awal ini berfungsi untuk memberikan informasi nilai *centroid* awal, variabel pada tabel ini meliputi :

- *Field* *id\_centroid\_awal* merupakan identitas *id\_centroid\_awal* pengelompokan siswa berdasarkan nilai akademik.
- *Field* *No Induk* merupakan nama *No induk* setiap siswa.
- *Field* *Kode\_Mapel* merupakan kode dari Mata Pelajaran.
- *Field* *C\_NH1-4, C\_UTS, C\_UAS* merupakan jarak *centroid* awal dari tiap-tiap data.

#### 3.6.4. Table Hasil *Centroid*

Tabel hasil *centroid* ini digunakan untuk menyimpan *id\_hasil\_centroid*, *Id\_Clsuter\_Awal*, *Kode\_Mapel*, *C\_NH1-4*, *C\_UTS*, dan *C\_UAS*, dan iterasi, dengan *id\_hasil\_centroid* sebagai *primary key*. Struktur dari tabel ini dapat dilihat pada tabel 3.16 :

**Tabel 3.16** Tabel Hasil Centroid

No	Nama	Tipe	Panjang	Keterangan
1	<i>Id_hasil_centroid</i>	Int	5	primary key
2	<i>Id_Cluster_Awal</i>	Int	10	
3	<i>Kode_Mapel</i>	Int	10	
4	<i>C_NH1</i>	Double	10	
5	<i>C_NH1</i>	Double	10	
6	<i>C_NH1</i>	Double	10	
7	<i>C_NH1</i>	Double	10	
8	<i>C_UTS</i>	Double	10	
9	<i>C_UAS</i>	Double	10	
10	<i>Iterasi</i>	Double	10	

Tabel hasil *centroid* ini berfungsi untuk memberikan informasi nilai *centroid* selama proses *Clustering*, variabel pada tabel ini meliputi :

- *Field* id\_hasil\_centroid merupakan identitas id\_hasil\_centroid pengelompokan siswa berdasarkan nilai akademik.
- *Field* Id\_Cluster\_Awal merupakan identitas dari id\_cluster awal
- *Field* Kode\_Mapel merupakan kode dari Mata Pelajaran.
- *Field* C\_NH1-4, C\_UTS, dan C\_UAS merupakan jarak centroid dari tiap-tiap data.
- *Field* Iterasi merupakan iterasi yang diikuti pada centroid dari tiap-tiap mata pelajaran.

### 3.6.5. Table Cluster Awal

Tabel *cluster* awal ini berfungsi untuk menyimpan id cluster awal, No Induk, id\_CentroidA, Kode Mapel, jarak terdekat, dan cluster, id cluster awal sebagai *primary key*. Struktur dari tabel ini dapat dilihat pada tabel 3.17 :

**Tabel 3.17** Tabel cluster awal

No	Nama	Tipe	Panjang	Keterangan
1	Id_cluster_awal	Int	5	primary key
2	No_Induk	Int	10	
3	Id_centroidA	Int	10	
4	Kode Mapel	Int	10	
5	Jarak Terdekat	Double	10	
6	Cluster	Int	5	

Tabel *cluster* awal ini berfungsi untuk memberikan informasi nilai *cluster* awal pada proses *Clustering*, variabel pada tabel ini meliputi :

- *Field* id\_cluster\_awal merupakan identitas id\_cluster\_awal dari proses *Clustering*.
- *Field* No\_Induk merupakan No\_induk dari setiap siswa.
- *Field* Id\_CentroidA merupakan identitas dari Id\_CentroidA
- *Field* Kode\_Mapel merupakan kode dari Mata Pelajaran.

- *Field* Jarak\_Terdekat merupakan jarak terdekat pada cluster tersebut.
- *Field* Cluster merupakan kluster yang diikuti dari setiap siswa.

### 3.6.6. Table Laporan Cluster

Tabel laporan *cluster* ini berfungsi untuk menyimpan id laporan cluster, No Induk, Id\_cluster\_Awal, Id\_Hasil\_Centroid, Kode Mapel, Jarak terdekat, *cluster*, dan iterasi, id laporan *cluster* sebagai *primary key*, Struktur dari tabel ini dapat dilihat pada tabel 3.18:

**Tabel 3.18** Tabel laporan cluster

No	Nama	Tipe	Panjang	Keterangan
1	Id_laporan_cluster	Int	5	primary key
2	No_Induk	Int	10	
3	Id_Cluster_Awal	Int	10	
4	Id_Hasil_Centroid	Int	10	
5	Kode_Mapel	Int	10	
6	Jarak_Terdekat	Double	10	
7	Cluster	Int	5	
8	Iterasi	Int	5	

Tabel laporan *cluster* ini berfungsi untuk memberikan informasi nilai *cluster* selama proses *Clustering*, variabel pada tabel ini meliputi :

- *Field* id\_laporan\_cluster merupakan identitas id\_laporan\_cluster dari proses *Clustering*.
- *Field* No\_Induk merupakan No\_induk dari setiap siswa.
- *Field* id\_cluster\_awal merupakan identitas dari cluster\_awal.
- *Field* id\_hasil\_centroid merupakan identitas dari hasil\_centroid.
- *Field* Kode\_Mapel merupakan kode dari Mata Pelajaran.
- *Field* Jarak\_Terdekat merupakan jarak terdekat pada *cluster* tersebut.
- *Field* Cluster merupakan kluster yang diikuti dari setiap siswa.

- *Field* Iterasi merupakan nilai iterasi yang sedang terjadi pada *cluster* tersebut.

### 3.6.7. Table Mata Pelajaran

Tabel mata pelajaran ini berfungsi untuk menyimpan kode mapel dan nama mata pelajaran, kode mapel sebagai *primary key*. Struktur dari tabel ini dapat dilihat pada tabel 3.19 :

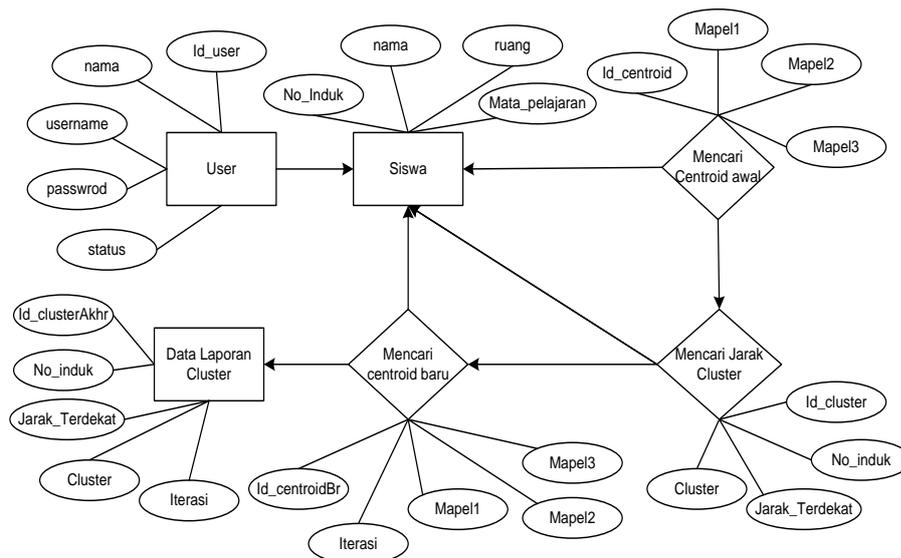
**Tabel 3.19** Tabel Mata Pelajaran

No	Nama	Tipe	Panjang	Keterangan
1	Kode_Mapel	Int	5	primary key
2	Mata Pelajaran	Varchar	25	

Tabel mata pelajaran ini berfungsi untuk memberikan informasi nama-nama mata pelajaran, variabel pada tabel ini meliputi :

- *Field* Kode\_Mapel merupakan identitas dari tiap-tiap Mata Pelajaran.
- *Field* Mata\_Pelajaran merupakan nama-nama Mata pelajaran.

### 3.7. ERD (Entity Relationship Diagram)

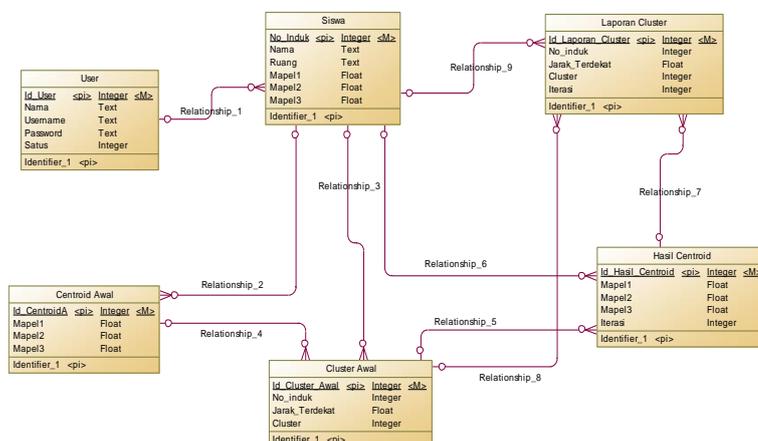


**Gambar 3.7** Entity Relationship Diagram

Keterangan :

- *User* merupakan data *store* dari tabel *User* yang berfungsi untuk memasukkan data siswa
- Siswa merupakan data *store* dari tabel Siswa yang berisi nilai-nilai siswa yang dibutuhkan dalam proses *Clustering*
- Mencari *Centroid* awal adalah proses dari tabel *Centroid\_Awal* yang mengambil data dari tabel Siswa dan hasilnya digunakan dalam mencari jarak terdekat.
- Mencari Jarak *Cluster* merupakan proses dari tabel *Cluster\_Awal* yang hasilnya digunakan dalam proses mencari *centroid* baru.
- Mencari *Centroid* baru merupakan proses dari tabel *Hasil Centroid* yang berisi data *centroid*.
- Data Laporan *cluster* merupakan data *store* dari tabel laporan *cluster* yang berisi nilai *cluster* dari proses *Clustering*.

### 3.8. CDM (Conceptual Data Model)



Gambar 3.8 Conceptual Data Model

Keterangan :

- Relationship\_1 : *user* memasukkan data siswa
- Relationship\_2 : data siswa diacak untuk memlih *centroid* awal
- Relationship\_3 : pengelompokan data siswa pada *cluster* awal

- Relationship\_4 : *centroid* awal digunakan membentuk *cluster* awal
- Relationship\_5 : data cluster awal digunakan untuk proses mencari *centroid* baru
- Relationship\_6 : data siswa digunakan untuk proses mencari *centroid* baru
- Relationship\_7 : hasil *centroid* dijadikan acuan untuk laporan *cluster*
- Relationship\_8 : cluster awal menjadi pembanding perubahan *cluster* untuk proses selanjutnya
- Relationship\_9 : pengelompokan data siswa sesuai *cluster*

### 3.9. Perancangan Antar Muka

Perancangan antar muka atau *interface* adalah bagian yang menghubungkan antara program dengan pemakai. *Interface* dari sistem dibuat dengan bahasa PHP berbasis *web*. Sistem ini berisikan informasi yang dikemas dalam beberapa menu.

#### 3.9.1 Halaman *Login*

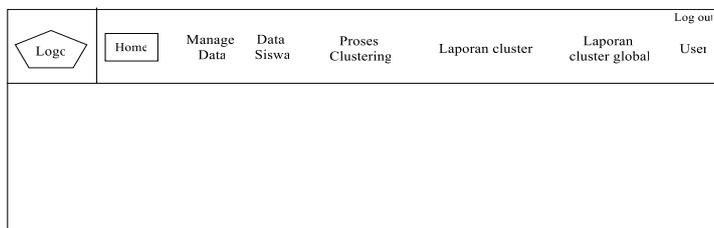
Halaman login ini digunakan admin dan *user* untuk masuk ke dalam sistem pengelompokan siswa. Tampilannya dapat dilihat pada gambar 3.9 :

The image shows a login form with the following elements:

- A logo consisting of a pentagon with the word "Logc" inside.
- The title: "Sistem Pengelompokan Siswa Berdasarkan Nilai Akademik dengan Metode K-Means".
- A label "Username :" followed by a text input field.
- A label "Password :" followed by a text input field.
- A button labeled "Log in" centered below the input fields.

**Gambar 3.9** *Interface* Halaman Login

### 3.9.2 Halaman Home Admin

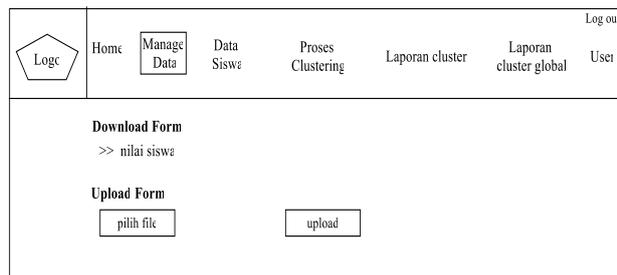


**Gambar 3.10** *Interface* Halaman Home Admin.

Keterangan Gambar 3.10 sebagai berikut :

1. *Header* : berisi logo instansi terkait dan menu.
2. *Menu* : berisi daftar pilihan menu yang disediakan oleh sistem.
3. *Viewer* : berisi tampilan dari menu yang telah dipilih, dalam menu *home* berisi latar belakang dan tujuan pembuatan sistem.

### 3.9.3 Halaman *Manage Data*



**Gambar 3.11** *Interface* Halaman *Manage Data*.

Keterangan Gambar 3.11 sebagai berikut :

1. *Header* : berisi logo instansi terkait dan menu.
2. *Menu* : berisi daftar pilihan menu yang disediakan oleh sistem.
3. *Viewer* : berisi tampilan dari menu yang telah dipilih, dalam menu *manage data* berisi *download form* yang bisa di *download* oleh admin dan kemudian di *upload* melalui *upload form*.



### 3.9.6 Halaman Laporan Cluster

**Gambar 3.14** Interface Halaman Laporan Cluster.

Keterangan Gambar 3.14 sebagai berikut :

1. *Header* : berisi logo instansi terkait dan menu.
2. *Menu* : berisi daftar pilihan menu yang disediakan oleh sistem.
3. *Viewer1* : berisi tampilan dari menu yang telah dipilih, dalam menu laporan berisi pilihan urutan berdasarkan yang dipilih.
4. *Viewer2* : berisi tampilan menu laporan *cluster* berisi laporan jumlah *cluster* data siswa beserta status *cluster*.
5. *Viewer3* : berisi tampilan detail tabel percobaan dari hasil proses *Clustering*.

### 3.9.7 Halaman Laporan Cluster Global

**Gambar 3.15** Interface Halaman Laporan Cluster Global.

Keterangan Gambar 3.15 sebagai berikut :

1. *Header* : berisi logo instansi terkait dan menu.
2. *Menu* : berisi daftar pilihan menu yang disediakan oleh sistem.

3. *Viewer1* : berisi tampilan dari menu yang telah dipilih, dalam menu laporan berisi pilihan urutan berdasarkan yang dipilih.
4. *Viewer2* : berisi tampilan dari menu yang telah dipilih, dalam menu laporan *cluster* berisi detail tabel *cluster* dari data siswa.

### 3.9.8 Halaman *User*

The screenshot shows a web application interface for managing users. At the top, there is a navigation menu with a 'Logc' logo and several menu items: Home, Manage Data, Data Siswa, Proses Clustering, Laporan cluster, Laporan cluster global, and Log out. Below the menu is a section titled 'DATA USER' containing a table with columns for No, Nama, Username, Password, and Aksi. The Aksi column contains 'Edit' and 'Delete' buttons. Below the table is a section titled 'TAMBAH USER' with input fields for Nama, Username, and Password, and 'Simpan' and 'Batal' buttons.

DATA USER					
No	Nama	Username	Password	Aksi	
				<input type="button" value="Edit"/>	<input type="button" value="Delete"/>

**TAMBAH USER**

Nama :

Username :

Password :

**Gambar 3.16** *Interface* Halaman Manage *User*.

Keterangan Gambar 3.16 sebagai berikut :

1. *Header* : berisi logo instansi terkait dan menu.
2. Menu : berisi daftar pilihan menu yang disediakan oleh sistem.
3. *Viewer1* : berisi tampilan dari menu yang telah dipilih, dalam menu manage *user* berisi data *user* yang telah ada pada tabel *user*.
4. *Viewer2* : berisi tampilan dari menu yang telah dipilih, dalam menu manage *user* berisi fasilitas untuk membuat *user* baru, simpan *user* dan batal.

### 3.10 Skenario Pengujian

Pengujian pengelompokan siswa berdasarkan nilai akademik dilakukan dengan percobaan perhitungan dengan *centroid* awal yang berbeda untuk mendapatkan nilai DBI terkecil.

1. Perhitungan dengan 5 kali percobaan.
2. Perhitungan dengan 10 kali percobaan.