

**PENGELOMPOKAN SISWA BERDASARKAN FAKTOR-FAKTOR
YANG MEMPENGARUHI KEBERHASILAN SISWA DALAM BELAJAR
MENGUNAKAN METODE CLUSTERING K-MEANS**

Yulya Muharmi¹

ABSTRACT

Grouping students based on factors that influence student success in learning aims to assist and facilitate the process of managing student data. In this study, the factors used are parental education, parental income, type of residence, a distance of a place to stay, transportation to school, student attitudes, interests, and the average value of students as a tool in determining the category of student success. Data Mining is a computer-based data processing to obtain a Knowledge. Data grouping students in one of the Data Mining method used is the method clustering with K-Means algorithm. RapidMiner application used to assist in processing the data to provide information for BK teachers and schools in the understanding of student data.

Keywords: *Data Mining, Clustering Algorithm K-Means, RapidMiner,*

INTISARI

Pengelompokan siswa berdasarkan faktor-faktor yang mempengaruhi keberhasilan siswa dalam belajar bertujuan untuk membantu dan mempermudah pada proses pengelolaan data siswa. Pada penelitian ini faktor-faktor yang digunakan adalah pendidikan orang tua, penghasilan orang tua, jenis tinggal, jarak tempat tinggal, alat transportasi ke sekolah, sikap siswa, minat serta nilai rata-rata siswa sebagai alat bantu dalam menentukan kategori keberhasilan siswa. Data Mining merupakan pengolahan data berbasis komputer untuk mendapatkan suatu Knowledge. Dalam pengelompokan data siswa salah satu metode Data Mining yang digunakan adalah metode Clustering dengan algoritma K-Means. Aplikasi RapidMiner digunakan untuk membantu dalam pengolahan data yang mampu memberikan informasi bagi guru BK dan pihak sekolah dalam pemahaman data siswa.

Kata Kunci : *Data Mining, Clustering, Algoritma K-Means, Rapidminer, Pengelompokan Siswa.*

¹ Dosen STMIK-AMIK Riau

PENDAHULUAN

Data Mining berisi pencarian *trend* atau pola yang diinginkan dalam *database* besar untuk membantu dalam pengambilan keputusan di waktu yang akan datang. Pola-pola ini dikenali oleh perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan berwawasan yang kemudian dapat dipelajari lebih teliti, yang mungkin saja menggunakan perangkat pendukung keputusan lainnya (Astuti, 2013). Ada banyak metode untuk melakukan *Data Mining*, antara lain *classification*, *regression*, *clustering* dan *association* (Sijabat, 2015).

Peran guru BK di SMA Al-Istiqamah dalam mengembangkan pribadi siswa masih ada yang belum efektif dan efisien. Salah satu adalah mengenai pemahaman terhadap data siswa. Pemahaman terhadap data siswa dilakukan untuk melihat keberhasilan siswa dalam belajar. Pengelompokan siswa ini dapat dilakukan dengan analisa *cluster*. *Cluster* pada data siswa sangat penting dalam mengambil langkah penempatan, pengarahan dan pengembangan siswa secara optimal

PENDEKATAN PEMECAHAN MASALAH

Knowledge Discovery in Database (KDD)

Data Mining adalah bagian integral dari penemuan pengetahuan dalam database (*Knowledge Discovery in Database/KDD*), yang merupakan proses keseluruhan mengubah data mentah menjadi informasi yang bermanfaat. Proses KDD ini terdiri dari serangkaian langkah-langkah transformasi, dari proses data *preprocessing* dan proses data *postprocessing* dari data yang merupakan hasil penggalian. Tujuan dari proses data *preprocessing* adalah untuk mengubah data input mentah menjadi format yang sesuai untuk analisis selanjutnya. Langkah-langkah yang dilakukan antara lain dengan

memperbaiki data yang kotor atau ganda, dan memilih catatan dan fitur yang relevan dengan proses pengelolaan data selanjutnya (Tan dkk, 2006 dalam Wahyudi dan Jananto, 2013).

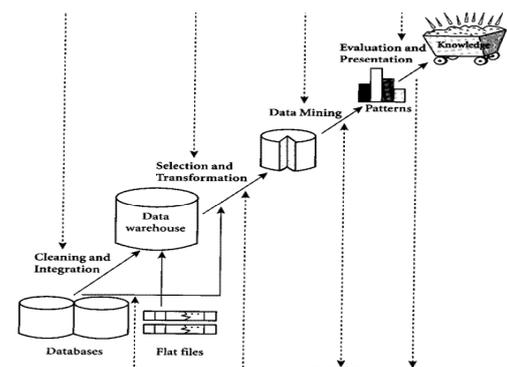
Data Mining

1. Definisi *Data Mining*

Data Mining (DM) adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara *manual*. Perkembangan *Data Mining* yang pesat tidak dapat lepas dari perkembangan teknologi informasi yang memungkinkan data dalam jumlah besar terakumulasi (Lindawati, 2008).

2. Tahapan *Data Mining*

Sebagai suatu rangkaian proses, *Data Mining* dapat dibagi menjadi beberapa tahap proses. Tahap-tahap tersebut bersifat *interaktif*, di mana pemakai terlibat langsung atau dengan perantara *knowledge base*. Tahapan tersebut dapat dilihat pada Gambar 2.1 berikut (Lindawati, 2008) :



Gambar 1. Tahap-Tahap *Data Mining*

Clustering

Clustering merupakan bagian dari ilmu *Data Mining* yang bersifat tanpa arahan (*unsupervised*). *Clustering* adalah proses pembagian data ke dalam kelas atau *cluster* berdasarkan tingkat kesamaannya. Dalam *clustering*, data yang memiliki

kesamaan dimasukkan ke dalam *cluster* yang sama, sedangkan data yang tidak memiliki kesamaan dimasukkan dalam *cluster* yang berbeda (Khotimah, 2014).

Algoritma K-Means

K-Means merupakan algoritma *clustering* yang berulang-ulang. Algoritma *K-Means* dimulai dengan pemilihan secara acak *K*, *K* di sini merupakan banyaknya *cluster* yang ingin dibentuk. Kemudian tetapkan nilai-nilai *K* secara random, untuk sementara nilai tersebut menjadi pusat dari *cluster* atau biasa disebut dengan *centroid*, *mean* atau "*means*".

Hitung jarak setiap data yang ada terhadap masing-masing *centroid* menggunakan rumus *Euclidian* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*.

Klasifikasikan setiap data berdasarkan kedekatannya dengan *centroid*. Lakukan langkah tersebut hingga nilai *centroid* tidak berubah (stabil) (Rismawan dan Kusumadewi,2008).

Faktor – Faktor yang Mempengaruhi Siswa Dalam Belajar

Faktor-faktor yang mempengaruhi belajar banyak jenisnya, tetapi dapat digolongkan menjadi dua golongan, yaitu faktor *intern* dan faktor *ekstern*. Faktor *intern* adalah faktor yang ada dalam diri individu yang sedang belajar, sedangkan faktor *ekstern* adalah faktor yang ada di luar individu (Slameto, 2013).

HASIL DAN PEMBAHASAN

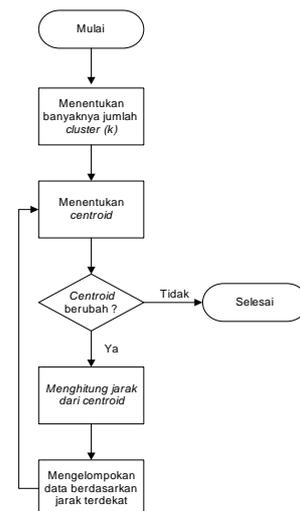
Dalam mengumpulkan data dilakukan dengan cara melakukam wawancara serta mengambil data kepada administrasi sekolah tersebut. Data yang digunakan adalah berdasarkan faktor-faktor yang mempengaruhi siswa dalam belajar menurut Slameto (2013). Faktor-faktor yang mempengaruhi keberhasilan belajar siswa banyak jenisnya, diantaranya faktor *Intern* dan faktor

Ekstern. Faktor *intern* adalah faktor yang ada dalam diri individu yang sedang belajar, sedangkan faktor *ekstern* adalah faktor yang ada di luar individu seperti yang sudah dijelaskan pada bab sebelumnya.

Analisa Clustering dengan Menggunakan Algoritma K-Means

Analisis *cluster* merupakan pengelompokan obyek-obyek data hanya berdasarkan pada informasi yang terdapat pada data, yang menjelaskan obyek dan relasinya.

Analisis *cluster* merupakan pengelompokan obyek-obyek data hanya berdasarkan pada informasi yang terdapat pada data, yang menjelaskan obyek dan relasinya. Tahapan algoritma *K-Means* dapat dilihat pada gambar dibawah:



Gambar 4.1 : Diagram Proses Flowchart K-Means

Penetapan Data

Sebelum melakukan proses pengelompokan data siswa, terlebih dahulu tentukan data yang akan dilakukan pengolahan agar tujuan penelitian dapat tercapai. Adapun langkah-langkahnya sebagai berikut :

1. Menentukan Atribut Yang Digunakan

Pada penelitian ini faktor keberhasilan yang digunakan dibatasi

hanya pada faktor keluarga dan faktor lingkungan yang terdiri dari :

- Latar belakang pendidikan ayah dan ibu
- Penghasilan ayah dan ibu
- Jenis tempat tinggal siswa seperti tinggal bersama orang tua, wali, kost, asrama, panti asuhan dan lainnya.
- Jarak tempat tinggal ke sekolah
- Alat transportasi ke sekolah seperti jalan kaki, kendaraan pribadi, kendaraan umum dan lainnya.
- Sikap siswa
- Minat

terbanyak muncul dengan mengurutkan frekuensi tertinggi ke terendah dan lakukan inisial data.

Tabel 2. Transformasi Pendidikan Ayah

No	Pendidikan Ayah	Frekuensi	Inisial
1	SMA Sederajat	58	1
2	SD Sederajat	41	2
3	SMP Sederajat	40	3
4	Putus SD	21	4
5	D4/S1	9	5
6	Tidak Sekolah	3	6
7	D3	2	7
8	S2	1	8

Tabel 1. Sampel Data Awal

No	Nama	Pendidikan Ayah	Pendidikan Ibu	Penghasilan Ayah	Penghasilan Ibu	Jarak Temp. Tinggal ke Sekolah	Alat Transportasi ke Sekolah	Jenis Tempat Tinggal	Minat
1	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
2	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
3	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
4	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
5	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
6	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
7	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
8	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
9	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
10	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
11	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
12	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
13	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
14	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
15	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
16	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
17	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
18	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
19	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
20	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
21	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
22	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
23	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
24	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
25	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
26	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
27	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
28	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
29	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
30	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
31	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
32	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
33	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
34	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
35	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
36	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
37	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
38	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
39	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
40	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
41	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
42	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
43	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
44	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
45	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
46	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
47	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
48	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
49	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
50	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
51	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
52	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
53	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
54	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
55	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
56	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
57	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
58	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
59	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
60	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
61	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
62	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
63	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
64	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
65	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
66	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
67	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
68	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
69	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
70	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
71	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
72	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
73	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
74	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
75	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
76	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
77	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
78	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
79	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
80	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
81	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
82	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
83	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
84	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
85	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
86	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
87	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
88	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
89	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
90	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
91	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
92	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
93	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
94	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
95	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
96	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
97	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
98	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
99	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
100	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000

2. Transformasi Data

Transformasi data dilakukan karena jenis data tidak berupa numeric maka data harus di tranformasikan terlebih dahulu dengan cara melakukan frekuensi pada data yang

Tabel 3. Sampel Data Siswa Proses Clustering

No	Nama	Pendidikan Ayah	Pendidikan Ibu	Penghasilan Ayah	Penghasilan Ibu	Jarak Temp. Tinggal ke Sekolah	Alat Transportasi ke Sekolah	Sikap	Minat
1	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
2	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
3	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
4	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
5	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
6	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
7	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
8	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
9	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000	1000	1000
10	Abdi Nur Hafidham	SD	SMP	1000	1000	1000	1000		

Cluster 2 : (86,83 ; 4 ; 4 ; 2 ; 1 ; 1 ; 1 ; 3 ; 1 ; 4 ; 1)

Menghitung jarak antara titik *centroid* dengan titik tiap objek dengan menggunakan *Euclidian Distance*.

3. Menghitung jarak dari *centroid*

Perhitungan jarak dari data ke-1 terhadap pusat *cluster* :

Centroid 1

$$= \sqrt{(86,17 - 78,61)^2 + (3 - 2)^2 + (1 - 1)^2 + (1 - 2)^2 + (3 - 2)^2 + (1 - 3)^2}$$

$$\sqrt{(1 - 1)^2 + (1 - 3)^2 + (1 - 1)^2 + (1 - 1)^2 + (1 - 7)^2} = 10,206$$

Centroid 2

$$= \sqrt{(78,61 - 78,61)^2 + (2 - 2)^2 + (1 - 1)^2 + (2 - 2)^2 + (2 - 2)^2 + (3 - 3)^2}$$

$$\sqrt{(1 - 1)^2 + (3 - 3)^2 + (1 - 1)^2 + (1 - 1)^2 + (7 - 7)^2} = 0,000$$

Centroid 3

$$= \sqrt{(81,06 - 78,61)^2 + (2 - 2)^2 + (1 - 1)^2 + (2 - 2)^2 + (1 - 2)^2 + (1 - 3)^2}$$

$$\sqrt{(1 - 1)^2 + (1 - 3)^2 + (1 - 1)^2 + (1 - 1)^2 + (7 - 7)^2} = 3,873$$

Centroid 4

$$= \sqrt{(78,83 - 78,61)^2 + (3 - 2)^2 + (4 - 1)^2 + (1 - 2)^2 + (1 - 2)^2 + (1 - 3)^2}$$

$$\sqrt{(1 - 1)^2 + (1 - 3)^2 + (1 - 1)^2 + (1 - 1)^2 + (10 - 7)^2} = 5,005$$

Centroid 5

$$= \sqrt{(78,44 - 78,61)^2 + (1 - 2)^2 + (2 - 1)^2 + (1 - 2)^2 + (2 - 2)^2 + (2 - 3)^2}$$

$$\sqrt{(1 - 1)^2 + (1 - 3)^2 + (1 - 1)^2 + (2 - 1)^2 + (10 - 7)^2} = 4,246$$

- Alokasikan masing-masing objek ke *centroid* terdekat.
- Lakukan iterasi, kemudian tentukan posisi *centroid* baru dengan cara menghitung rata-rata dari data-data yang berada pada *centroid* yang sama.
- Lakukan perulangan pada langkah ke-3 hingga posisi data tidak mengalami perubahan.

Iterasi-1					
Cluster 0	Cluster 1	Cluster 2	Cluster 0	Cluster 1	Cluster 2
9,573	5,828	3,730	0	0	1
2,744	14,080	8,844	1	0	0
2,192	12,184	7,344	1	0	0
4,046	15,133	10,534	1	0	0
3,038	15,791	10,832	1	0	0
3,033	14,669	10,142	1	0	0
4,004	16,249	11,689	1	0	0
3,805	15,620	11,116	1	0	0
12,844	3,366	7,268	0	1	0
10,702	5,435	4,632	0	0	1
9,601	6,207	4,168	0	0	1
14,661	3,366	8,413	0	1	0
7,112	9,343	2,636	0	0	1
7,494	9,403	3,603	0	0	1
9,750	5,500	3,661	0	0	1
7,233	10,263	4,578	0	0	1
7,526	9,618	3,717	0	0	1
9,465	10,727	5,420	0	0	1
9,932	9,815	5,329	0	0	1
6,848	11,443	4,444	0	0	1

Tabel 4. Hasil Perhitungan Jarak dan Pengelompokan Data Iterasi ke-0

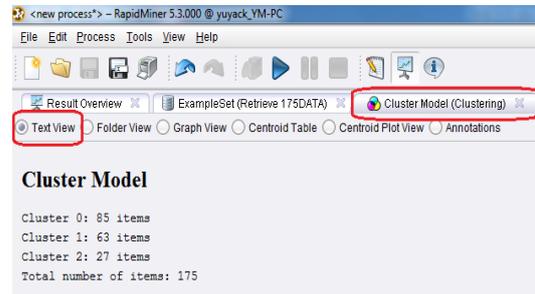
Iterasi-0					
Cluster 0	Cluster 1	Cluster 2	Cluster 0	Cluster 1	Cluster 2
10,206	5,595	5,333	0	0	1
0,000	13,355	11,427	1	0	0
3,873	11,659	9,762	1	0	0
5,005	14,841	12,490	1	0	0
4,246	15,178	13,244	1	0	0
4,758	14,177	12,656	1	0	0
5,104	15,788	14,036	1	0	0
5,511	15,311	13,073	1	0	0
13,355	0,000	5,679	0	1	0
11,427	5,679	0,000	0	0	1
10,673	6,316	5,947	0	0	1
15,518	6,732	7,034	0	1	0
7,730	8,901	6,664	0	0	1
7,752	9,295	6,922	0	0	1
10,677	4,779	4,364	0	0	0
7,441	9,394	7,280	1	0	1
7,890	9,196	7,299	0	0	1
10,122	11,720	7,895	0	0	1
10,694	10,920	7,746	0	0	1
7,220	11,119	8,323	1	0	0

Tabel 5. Hasil Perhitungan Jarak dan Pengelompokan Data Iterasi ke-1

Tabel 6. Hasil Perhitungan Jarak dan Pengelompokan Data Iterasi ke-2

Iterasi-2					
Cluster 0	Cluster 1	Cluster 2	Cluster 0	Cluster 1	Cluster 2
11,139	5,828	3,946	0	0	1
3,488	14,080	8,401	1	0	0
3,372	12,184	7,000	1	0	0
3,082	15,133	10,311	1	0	0
1,811	15,791	10,471	1	0	0
1,994	14,669	9,865	1	0	0
2,549	16,249	11,356	1	0	0
2,273	15,620	10,855	1	0	0
14,130	3,366	7,701	0	1	0
12,150	5,435	5,049	0	0	1
11,022	6,207	4,528	0	0	1
15,827	3,366	9,085	0	1	0
8,954	9,343	2,055	0	0	1
9,251	9,403	3,205	0	0	1
11,297	5,500	3,961	0	0	1
9,080	10,263	3,958	0	0	1
9,325	9,618	3,238	0	0	1
10,921	10,727	5,441	0	0	1
11,327	9,815	5,521	0	0	1
8,687	11,443	3,830	0	0	1

Pada *Cluster Model* (*Clustering*) dapat dilihat jumlah dari ke-3 *cluster*.



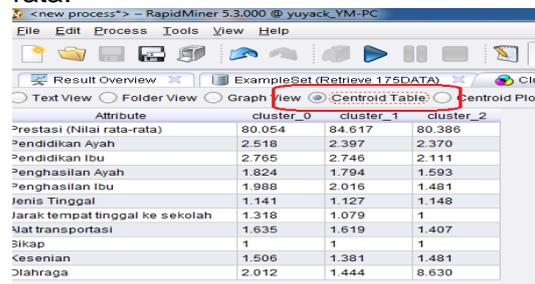
Gambar 2. Tampilan *Cluster Model* (*Text View*)

Pada iterasi ke-1 dan iterasi ke-2 hasil *cluster* tidak berubah maka perulangan tidak dilakukan lagi. Dapat ditarik kesimpulan bahwa *cluster 0* terdiri dari 7 siswa, *cluster 1* terdiri dari 2 siswa dan *cluster 2* terdiri dari 11 siswa terlihat pada tabel 7.

Tabel 7. Hasil Akhir *Cluster*

Nama	Prestasi (Nilai Rata-Rata)	Pendidikan Ayah	Pendidikan Ibu	Penghasilan Ayah	Penghasilan Ibu	Jenis Tameng	Jarak Tempat Tinggal Ke Sekolah	Alat Transportasi	Sikap	Kesenian	Mind Outrage
Cluster 0											
Berprestasi Kategori Rendah											
Gelis Satria	78,61	2	1	2	2	3	1	3	1	1	7
Andrya Zuhairi	81,06	2	1	2	1	1	1	1	1	1	7
Kharul Ramadhani	78,83	3	4	1	1	1	1	3	1	1	10
Lentis Wastilla	78,44	1	2	1	2	2	1	1	1	2	10
M. Diky Alfians	79,89	1	2	1	3	1	1	1	1	1	10
M. Rafiq	78,83	1	1	3	2	3	1	1	1	1	11
Mari Agengiti	79,22	2	3	2	1	1	1	1	1	2	11
Cluster 1											
Berprestasi Kategori Tinggi											
Estiani	90,33	2	3	1	2	3	1	3	1	1	1
Hilmi Rizki Aqyan	90,89	7	5	3	4	1	1	1	1	1	1
Cluster 2											
Berprestasi Kategori Sedang											
Caesarendra	86,63	4	4	2	1	1	1	3	1	4	1
Helya Andara	86,22	1	5	3	4	1	2	1	1	1	2
Danu Tedyani	86,17	3	1	1	3	1	1	1	1	1	1
Jadi Johanda	81,89	2	3	1	2	1	1	1	1	1	1
Abdi Muthalib	81,94	4	1	1	2	1	1	1	1	1	1
Melawati	86,61	2	3	1	1	1	1	1	1	1	1
Almal Yusri Fauzi	81,83	1	1	3	1	2	1	1	1	2	1
Syihan Asif	82,11	2	1	1	4	1	1	1	1	2	1
Andryoga Ananda	80,72	5	7	3	4	1	1	1	1	3	1
Andri Restu	81,83	5	8	3	3	1	1	1	1	1	1
Azzura	79,67	2	3	1	3	1	1	1	1	2	1

Pada *centroid table* dapat kita lihat bahwa *cluster 0* merupakan kelompok siswa yang berprestasi kategori rendah, *cluster 1* merupakan kelompok siswa yang berprestasi kategori tinggi, dan *cluster 2* merupakan kelompok siswa yang berprestasi kategori sedang. Hal ini dapat dibaca pada *centroid* masing-masing *cluster* pada *attribute* nilai rata-rata.



Gambar 3. Tampilan *Centroid Table*

Hasil Implementasi

Setelah dilakukan analisis data seperti penjelasan pada bab sebelumnya maka langkah selanjutnya adalah melakukan implementasi. Implementasi dilakukan dengan menggunakan aplikasi *RapidMiner 5.3* yang akan membantu mempermudah dalam pengelompokan data siswa pada SMA Al – Istiqamah Simpang Empat Pasaman Barat yang akan membantu dalam proses pengambilan keputusan.

KESIMPULAN

Dari pembahasan yang telah dikemukakan dapat diambil kesimpulan bahwa :

1. Metode *Clustering K-Means* dapat membantu mengelompokkan data siswa berdasarkan faktor-faktor yang mempengaruhi keberhasilan siswa dalam belajar sehingga membantu dalam penempatan, pengembangan dan pengarahannya terhadap siswa.
2. Didapatkan 3 kategori prestasi siswa, yaitu berprestasi kategori rendah, berprestasi kategori sedang, dan

- berprestasi kategori tinggi dimana masing-masing cluster memiliki nilai rata-rata yang berbeda.
3. Di SMA Al-Istiqamah Simpang Empat Pasaman Barat siswa yang berprestasi kategori tinggi jumlah siswanya sedang, dengan penghasilan orang tuanya kurang mampu, domisili siswa kurang dari 1 km, tinggal bersama orang tua, dengan pendidikan orang tua rendah.

DAFTAR PUSTAKA

- [1] Arora, Badal. 2013. Evaluating Student's Performance Using K-Means Clustering. *International Journal of Computer Science And Technology*. Volume 4, Issue 2, April 2013 . ISSN 0976-8491.
- [2] Dewi, Utomo, Yulianto. 2013. Identification of Potential Student Academic Ability using Comparison Algorithm K-Means and Farthest First. *International Journal of Computer Applications*. Volume 63– No.17, February 2013.
- [3] Ediyanto, Mara dan Satyahadewi. 2013. Pengklasifikasian Karakteristik dengan Metode K-Means Cluster Analysis. *Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster)*. Volume 02 , No. 2.
- [4] Fadlina. 2014. Data Mining Untuk Analisa Tingkat Kejahatan Jalanan Dengan Algoritma Association Rule Metode Apriori. *Informasi dan Teknologi Ilmiah*. Volume : III, Nomor : 1, Mei 2014. ISSN : 2339-210X.
- [5] Giri Sucahyo, Y. 2003. Penerapan Data Mining. *Artikel Populer Ilmu Komputer*.
- [6] Khotimah, T. 2014. Pengelompokan Surat Dalam Al Qur'an Menggunakan Algoritma K-Means. *Jurnal SIMETRIS*. Vol 5 No 1 April 2014. ISSN: 2252- 4983.
- [7] Lindawati. 2008. Data Mining dengan Teknik Clustering dalam Pengklasifikasian Data Mahasiswa." *Jurnal Informatika*. Mei 2008, ISSN : 1979-2328.
- [8] Saleh, Minhayati. 2008. Pengaruh Motivasi, Faktor Keluarga, Lingkungan Kampus Dan Aktif Berorganisasi Terhadap Prestasi Akademik. *Jurnal Phenomenon*, Volume 4 Nomor 2, Oktober 2014
- [9] Nur Wahyudi, E dan Jananto,A. 2013. Final Report Penilaian Kinerja Dosen oleh Mahasiswa pada Satu Periode Tahun Akademik menggunakan Teknik Klustering. *Jurnal Teknologi Informasi DINAMIK*. Volume 18, No.2, Juli 2013, ISSN : 0854-9524.
- [10] Nur Azizah, Prasetyo dan Arijanto. 2014. Usulan Peningkatan Prestasi Akademik Berdasarkan Faktor-Faktor yang Mempengaruhinya. *Jurnal Online Institut Teknologi Nasional*. Vol.02, No.04, Oktober 2014. ISSN: 2338-5081
- [11] Oscar Ong, J. 2013. Implementasi Algoritma K-Means Clustering Untuk menentukan strategi marketing President university. *Jurnal Ilmiah Teknik Industri*, Vol. 12, No. 1, Juni 2013. ISSN 1412-6869.
- [12] Oyelade, Oladipupo, dan Obagbuwa. 2010. Application of K-Means Clustering Algorithm for Prediction of Students' Academic Performance. *International Journal of Computer Science and Information Security*. Vol. 7, No 1, 2010. ISSN 1947-5500.
- [13] Rismawan, T dan Kusumadewi, S. 2008. Aplikasi K-Means Untuk Pengelompokan Mahasiswa Berdasarkan Nilai Body Mass Index (Bmi) & Ukuran Kerangka. *Seminar Nasional Aplikasi*

- Teknologi Informasi. ISSN: 1907-5022.
- [14] Sijabat, Alimancon. 2015. Penerapan Data Mining untuk Pengolahan Data Siswa dengan Menggunakan Metode Decision Tree. Jurnal Informasi dan Teknologi Ilmiah. Volume 5 No 3. Februari 2015, ISSN : 2339-210X.
- [15] Shovon, Haque. 2012. Prediction of Student Academic Performance by an Application of K-Means Clustering Algorithm. International Journal of Advanced Research in Computer Science and Software Engineering. Volume 2, Issue 7, July 2012. ISSN : 2277 128X.