

Implementasi *k-Means Clustering* untuk Analisis Nilai Akademik Siswa Berdasarkan Nilai Pengetahuan dan Keterampilan

Diwa Oktario Dacwanda ¹⁾, Yessica Nataliani ²⁾
Program Studi Sistem Informasi
Fakultas Teknologi Informasi
Universitas Kristen Satya Wacana Salatiga
Jl. O. Notohamidjojo 1-10, Salatiga 50711, Indonesia
Email: ¹⁾682017036@student.uksw.edu, ²⁾yessica.nataliani@uksw.edu

Received: 24-06-2021

Riwayat artikel:
Revised: 11-08-2021

Accepted: 16-08-2021

Abstract

*The level of student success is a reflection of the quality of the world of education. The world of education needs to improve the quality of students and student performance through facilities, infrastructure, and human resources. Information systems are now a resource that can be used to improve the quality and achievement of students, by analyzing student achievement. This analysis is carried out by classifying students' achievements from their ability to master the subject matter. Data were obtained from 266 students and 12 subjects with two categories of assessment, namely knowledge and skills according to the 2013 curriculum. This study used the method *k-means clustering*, where the determination of the number of clusters was determined by the validity index Davies Bouldin. Three groups of achievements are generated based on academic scores, namely smart, moderate, and sufficient. As a comparison, the grouping of student achievements is also done by looking at the value of knowledge alone. As many as 10.15 percent of all students experienced a shift cluster. Student achievement groupings can be used to help schools determine the appropriate teaching system.*

Keywords: *student achievements, clustering, k-means, knowledge, skill*

Abstrak

Tingkat keberhasilan siswa merupakan cerminan dari kualitas dunia pendidikan. Dunia pendidikan perlu meningkatkan kualitas siswa dan kinerja siswa melalui sarana, prasarana, dan sumber daya manusia. Sistem informasi kini menjadi sumber daya yang dapat digunakan untuk meningkatkan kualitas dan prestasi siswa, dengan melakukan analisis prestasi siswa. Analisis ini dilakukan dengan mengelompokkan prestasi siswa dari kemampuannya dalam menguasai materi pelajaran. Data didapat dari 266 siswa dan 12 mata pelajaran dengan dua kategori penilaian yaitu pengetahuan dan keterampilan sesuai dengan kurikulum 2013. Penelitian ini menggunakan metode *k-means clustering*, dimana penentuan jumlah *cluster*-nya ditentukan dengan indeks validitas Davies Bouldin. Tiga kelompok prestasi yang dihasilkan berdasarkan nilai akademik yaitu pintar, sedang, dan cukup. Sebagai pembandingan, pengelompokan prestasi siswa ini juga dilakukan dengan melihat nilai pengetahuan saja. Sebanyak 10.15 persen dari keseluruhan siswa mengalami

perpindahan cluster. Pengelompokan prestasi siswa dapat digunakan dalam membantu sekolah menentukan sistem pengajaran yang tepat.

Kata kunci: prestasi akademik siswa, *clustering*, *k-means*, pengetahuan, keterampilan

1. Pendahuluan

Tidak sedikit upaya yang dilakukan pihak sekolah maupun pihak di bidang pendidikan untuk meningkatkan prestasi siswa di bidang akademik yang berguna untuk mencapai standar pendidikan nasional. Tingkat keberhasilan siswa merupakan cerminan dari kualitas dunia pendidikan. Dunia pendidikan perlu meningkatkan kualitas dan kinerja siswa melalui sarana, prasarana, dan sumber daya manusia. Tidak hanya itu, sistem informasi kini menjadi sumber daya yang dapat digunakan untuk meningkatkan kualitas dan prestasi siswa. Teknologi *data mining* merupakan sebuah teknologi pemrosesan data yang dapat menemukan dan mengidentifikasi pola data yang tersembunyi. Hasil pengolahan data tersebut berguna dalam pengambilan keputusan ke depannya [1].

Dengan adanya teknologi *data mining*, dapat dilakukan analisis prestasi siswa. Analisis ini dilakukan dengan mengelompokkan siswa dari kemampuannya dalam memahami materi pelajaran. Pembagian kelompok belajar menggunakan proses *clustering* dilakukan dengan membagi sekelompok siswa menjadi subset yang disebut *cluster*. *Clustering* adalah suatu metode dalam *data mining* yang berguna dalam menganalisis data agar lebih akurat dalam memecahkan masalah pengelompokan data atau membagi kumpulan data menjadi beberapa himpunan bagian. Tujuan dari *clustering* yaitu untuk menetapkan data ke dalam suatu kelompok sehingga hubungan antar anggota pada *cluster* yang sama menjadi kuat, sementara hubungan antar anggota pada *cluster* yang berbeda menjadi lemah [2]. Objek dalam suatu *cluster* memiliki kemiripan karakteristik tetapi memiliki karakteristik yang berbeda dengan objek pada *cluster* lainnya. Oleh karenanya, *clustering* sangat berguna dalam menetapkan kelompok atau *cluster* yang tidak diketahui pada data.

Dalam melakukan penelitian ini digunakan metode *k-means clustering* untuk mengelompokkan kemampuan akademik siswa. Dengan teknik *clustering k-means* diharapkan pihak sekolah maupun pihak di bidang pendidikan dapat mendata siswa dengan kemampuannya masing-masing dan dapat dilakukan pengajaran dengan metode yang tepat sehingga dapat meningkatkan kualitas siswa dan prestasi akademik siswa. Fokus penelitian ini pada pengelompokan prestasi siswa di SMP Negeri 2 Salatiga dengan metode *clustering k-means*, dimana siswa dikelompokkan menjadi tiga *cluster* yaitu pintar, sedang, dan cukup, berdasarkan nilai akademiknya, yang terdiri dari dua kategori, yaitu kategori pengetahuan dan kategori keterampilan sesuai dengan kurikulum 2013. Selanjutnya, dalam penelitian ini akan dibandingkan hasil pengelompokan menggunakan gabungan

nilai pengetahuan dan keterampilan dan hasil pengelompokan dengan nilai pengetahuan saja, untuk dilihat apakah nilai keterampilan memang mempengaruhi nilai akademik siswa

2. Kajian Pustaka

Adapun penelitian terdahulu yang membahas pengelompokan *k-means* di antaranya adalah penggunaan *k-means* untuk mengelompokkan nilai UTS siswa, yang menghasilkan nilai UTS tinggi, sedang, dan rendah [3]. Selain itu, algoritma *k-means* digunakan untuk pengelompokan prestasi siswa tinggi, menengah, cukup di SMP Negeri I Sukahening [4]. Algoritma *k-means* juga digunakan untuk mendukung Sistem Informasi Akademik dan menentukan kelas unggulan. Dengan algoritma *k-means*, sistem mengolah nilai siswa dan mengelompokkan sesuai dengan nilai terdekat dengan titik pusat *cluster* [5]. Metode *k-means clustering* juga dimanfaatkan dalam memutuskan penjurusan siswa SMA [6]. Selain itu, *k-means* juga digunakan untuk menentukan potensi produksi buah-buahan di Daerah Istimewa Yogyakarta. Penelitian tersebut mendeskripsikan penerapan pengelompokan daerah penghasil buah dengan *k-means clustering* [7].

k-Means adalah algoritma yang membagi data menjadi beberapa kelompok berbeda. Data yang memiliki kesamaan kemiripan dikelompokkan menjadi satu *cluster* dan data yang memiliki perbedaan dikelompokkan ke kelompok yang berbeda [8]. Algoritma ini bekerja dengan meminimalkan jarak antara data dan pusat *cluster*. Dalam proses *clustering*, algoritma ini bergantung pada data yang diperoleh dalam proses tersebut dan kesimpulan yang akan diambil.

Salah satu menentukan jumlah *cluster* pada k-means adalah dengan indeks *Davies-Bouldin* (DB). Indeks DB adalah suatu metode yang berguna untuk menghitung validitas *cluster*, dengan cara memaksimalkan jarak data antar *cluster* (*between-cluster*) dan meminimalkan jarak data dalam *cluster* (*within-cluster*). Tahapan perhitungan indeks DB adalah sebagai berikut [9].

1. Menghitung *Sum of Square Within-cluster* dengan Rumus (1).

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \quad (1)$$

2. Menghitung *Sum of Square Between-cluster* dengan Rumus (2).

$$SSB_{i,j} = d(c_i, c_j) \quad (2)$$

3. Menghitung rasio dengan Rumus (3).

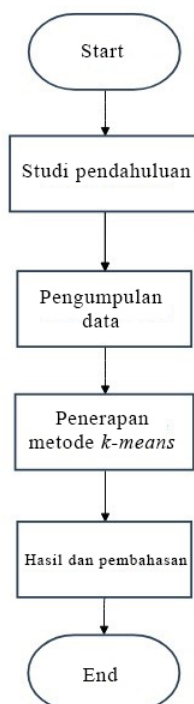
$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (3)$$

4. Menghitung indeks DB dengan Rumus (4).

$$\text{Indeks DB} = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (4)$$

3. Metode Penelitian

Penelitian ini menggunakan metode kualitatif. *Clustering k-means* yang diterapkan dalam pembentukan *cluster* siswa berdasarkan nilai dari data siswa. Penelitian dilaksanakan di SMP Negeri 2 Salatiga, yang menerapkan kurikulum 2013. Kurikulum 2013 adalah kurikulum lanjutan dari kurikulum berbasis kompetensi yang diterapkan pada tahun 2004. Kurikulum 2013 mencakup penilaian sikap, kemampuan, dan keterampilan siswa [10]. Pengelompokan siswa dilihat dari nilai akademik berdasarkan nilai pengetahuan dan keterampilan menggunakan algoritma *k-means*. Jumlah *cluster* siswa ditentukan menggunakan indeks *Davies-Bouldin*.



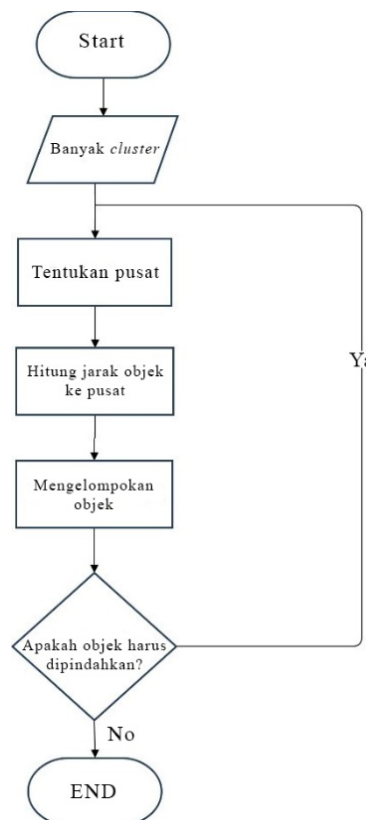
Gambar 1 *Flowchart* Penelitian

Dari Gambar 1 dapat dijelaskan tahapan penelitian sebagai berikut.

1. Studi pendahuluan, yang merupakan langkah awal penelitian dengan mengumpulkan dan meneliti data yang berkaitan sehingga diperoleh informasi tentang penelitian yang akan dilakukan.
2. Pengumpulan data, dimana teknik pengumpulan dan pengolahan data dalam penelitian ini dilakukan dalam berbagai bentuk, yaitu melalui observasi dan wawancara. Dalam penelitian ini, pihak yang dijadikan sebagai narasumber adalah pihak yang dianggap memiliki informasi yang dibutuhkan dalam penelitian. Teknik pengumpulan data juga digunakan dalam bentuk studi pustaka, yaitu pengumpulan data, informasi, dan teori-teori terkait dari literatur untuk mendukung analisis penelitian.

3. Penerapan metode *k-Means*, dimana data yang dimiliki akan diolah dengan algoritma *k-means* dan didapatkan hasil pengelompokan siswa berdasarkan tiga kategori siswa berprestasi tinggi, sedang, cukup, yang didapat dari indeks *Davies-Bouldin*. Setelah melakukan alur penelitian pada Gambar 1, maka selanjutnya data diproses agar membentuk tiga kategori siswa yang berprestasi tinggi, sedang, cukup. Data tersebut akan diproses menggunakan algoritma *k-means*.

Secara umum langkah yang digunakan dalam pengelompokan dengan metode *k-means* dapat dilihat pada Gambar 2 dan dijelaskan sebagai berikut [11].



Gambar 2 Flowchart Algoritma *k-Means*

1. Tentukan *k* sebagai banyaknya *cluster*, kemudian tentukan *centroid* awal, dimana *centroid* awal ditentukan secara acak berdasarkan data yang tersedia sebanyak *k cluster*.
2. Hitung jarak dari data ke-*i* ke *centroid* ke-*c* dengan persamaan *Euclidean* seperti pada Rumus (5).

$$d(x_i, v_c) = \|x_i - v_c\| = \sqrt{\sum_{j=1}^m (x_{ij} - v_{cj})^2} \quad (5)$$

dimana:

$d(x_i, v_c)$: jarak dari data ke-*i* ke *centroid* ke-*c*, $i = 1, 2, \dots, n$; $c = 1, 2, \dots, k$.

x_{ij} : data ke-*i* dan atribut ke-*j*, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$.

v_{cj} : *centroid* ke- c dan atribut ke- j , $c = 1, 2, \dots, k$; $j = 1, 2, \dots, m$.

3. Alokasikan setiap data ke dalam *cluster* yang mempunyai jarak terdekat.

4. Hitung *centroid* baru menggunakan Rumus (6).

$$v_{cj} = \frac{\sum_{x_i \in \text{Clust}_c} x_{ij}}{|\text{Clust}_c|} \quad (6)$$

dimana:

$x_i \in \text{Clust}_c$: data ke- i pada *cluster* ke- c , $i = 1, 2, \dots, n$; $c = 1, 2, \dots, k$.

$|\text{Clust}_c|$: jumlah data pada *cluster* ke- c , $c = 1, 2, \dots, k$.

x_{ij} : data ke- i dan atribut ke- j , $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$.

v_{cj} : *centroid* ke- c dan atribut ke- j , $c = 1, 2, \dots, k$; $j = 1, 2, \dots, m$.

5. Hitung selisih *centroid* baru dengan *centroid* lama. Jika tidak ada perubahan signifikan, maka selesai. Jika tidak, maka ulangi langkah 2.

Data mining merupakan bagian dari seri *Knowledge Discovery in Databases* (KDD). KDD mencakup eksplorasi ilmiah, interpretasi, dan visualisasi model kumpulan data dengan teknik terintegrasi. Rangkaian langkah pengolahan untuk pengelompokan siswa dapat dijelaskan sebagai berikut [12].

1. *Data selection*

Pemilihan data berguna untuk menganalisis data terkait untuk perhitungan *clustering*, karena terkadang semua data tidak diperlukan dalam proses *data mining*. Proses pemilihan data dilakukan sebelum proses *k-means clustering* dilakukan. 24 atribut dipilih dari 30 atribut data awal dengan menghilangkan atribut Kelas, NISN, Nama, Pramuka, Eks_2, Eks_3, karena atribut-atribut tersebut tidak dibutuhkan dalam proses *clustering*. 24 atribut yang digunakan merupakan nilai 12 mata pelajaran, dimana setiap mata pelajaran memiliki dua komponen nilai, yaitu nilai pengetahuan dan nilai keterampilan.

2. *Data cleaning*

Setelah tahap pengumpulan dan penyaringan data, maka dilakukan proses *data cleaning*, yaitu proses yang dilakukan dengan cara melakukan penghapusan data duplikasi, penghapusan *noise*, dan penghapusan data yang tidak konsisten. Karena tidak ada data yang perlu dibersihkan, maka tahap ini tidak dilakukan.

3. *Data transformation*

Data transformation dilakukan untuk menyesuaikan format data mentah menjadi format yang sesuai dengan proses *data mining* agar dapat terbaca dan dapat dilakukan proses perhitungan. Dalam penelitian ini data akademik siswa disusun dengan format .xlsx yang nantinya akan digunakan dalam pengolahan data menggunakan *software* RapidMiner.

4. Data mining

Data mining merupakan proses menggunakan metode atau algoritma tertentu untuk menemukan keunikan data. Banyak metode atau algoritma yang bisa digunakan dalam data mining. Oleh karena itu, pemilihan metode *data mining* perlu memperhatikan tujuan dan proses KDD secara keseluruhan.

5. Interpretation/evaluation

Model informasi yang dihasilkan perlu disajikan dalam mode tampilan yang dipahami oleh pihak-pihak terkait. Tahap ini dilakukan dalam rangka mencari pengetahuan, termasuk memeriksa apakah pola atau informasi yang didapatkan bertentangan atau sudah sesuai dengan fakta atau asumsi dari data yang sudah ada sebelumnya.

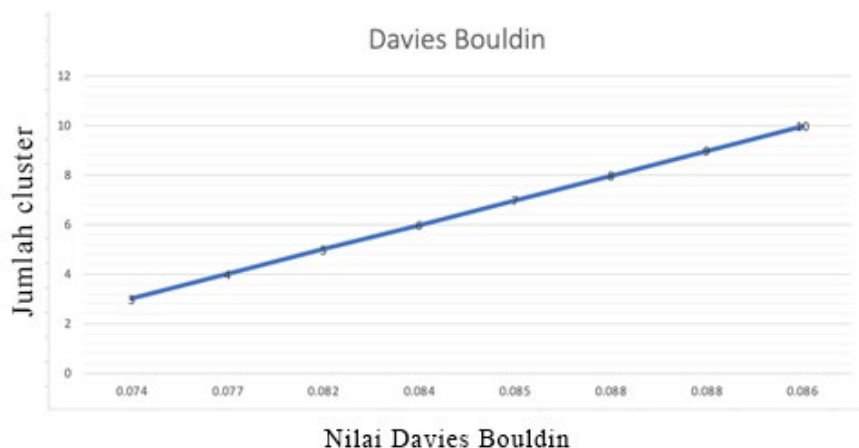
4. Hasil dan Pembahasan

Sumber data diambil dari nilai akademik siswa kelas IXA - IXH SMP Negeri 2 Salatiga. Data tersebut merupakan data nilai akademik dengan kurikulum 2013 yang memiliki dua kategori nilai yaitu nilai pengetahuan (Peng) dan keterampilan (Ketr). Data akademik yang diambil dalam penelitian ini adalah data nilai satu semester, dengan 12 mata pelajaran, yaitu Pendidikan Agama dan Budi Pekerti (PABK), Pendidikan Pancasila dan Kewarganegaraan (PPKn), Bahasa Indonesia (INDO), Matematika (MAT), Ilmu Pengetahuan Alam (IPA), Ilmu Pengetahuan Sosial (IPS), Bahasa Inggris (ING), Seni Budaya (SENI), Pendidikan Jasmani, Olahraga dan Kesehatan (PENJAS), Prakarya (PRAK), Bahasa Jawa (JAWA), dan TIK-BK. Gambar 3 merupakan data akademik siswa, yang terdiri dari 266 siswa dan 24 atribut penilaian mata pelajaran.

Nama	Peng_v	Ketr_f	Peng_f	Ketr_iv	Peng_v	Ketr_iv	Peng_v	Ketr_iv	Peng_v	Ketr_iv	Peng_v	Ketr_j	Peng_v	Ketr_pi	Peng_v	Ketr_f	Peng_j	Ketr_j	Peng_t	Ketr_t				
AALIYAH	90	93	90	95	85	85	86	95	88	86	86	93	94	89	95	89	84	86	91	90	91	92	92	89
AIDYA B.	94	95	90	94	83	86	85	95	92	87	89	92	91	89	91	91	84	85	91	89	92	90	91	80
AINAYA	89	93	92	94	82	82	86	95	91	87	89	89	89	89	95	94	90	87	91	89	90	90	91	89
ALFI KUN	95	95	94	95	87	87	89	95	93	92	89	92	92	93	95	92	89	87	95	91	94	92	94	90
ALIA NUI	92	89	83	94	83	80	80	92	81	87	86	94	90	88	85	83	83	86	85	90	93	92	87	80
ALVITA F	89	93	95	95	91	85	88	95	91	90	86	92	90	89	87	91	91	87	94	87	96	92	94	90
ANGELIK	84	93	88	93	85	83	85	94	82	88	87	93	87	87	93	90	88	86	90	89	91	89	94	90
ANGGRA	97	99	92	94	87	84	85	95	88	88	86	92	89	92	93	83	86	86	93	90	95	92	94	90
ANISHA I	85	85	88	95	84	80	81	94	83	83	85	94	87	85	87	85	82	85	88	89	88	90	91	80
ARNESTI	91	92	92	95	82	83	85	92	90	85	88	93	91	88	94	91	91	84	92	87	89	92	89	89
AULIA PI	89	92	87	94	83	85	84	94	92	90	89	89	93	87	95	92	90	86	91	89	91	89	92	89
BRIAN F	90	92	91	94	84	80	83	93	83	88	86	94	87	83	88	95	85	83	93	91	91	87	88	89
DEA JUM	94	97	85	94	84	83	80	93	82	85	87	92	85	90	84	88	83	85	95	91	89	90	95	89
DELVI AF	86	93	89	95	83	84	92	92	88	87	86	93	85	89	90	90	84	86	89	90	90	92	89	89
FABHIAN	91	92	89	95	84	82	85	93	90	85	88	94	92	90	92	90	84	84	93	89	89	92	91	89
HANAN I	91	92	94	95	84	84	85	94	87	87	88	93	87	90	94	93	89	86	93	90	90	89	91	89
JASMINE	95	95	92	95	84	84	88	95	89	91	91	91	93	88	97	92	88	89	95	90	92	92	94	89
JELITA R.	95	93	94	95	85	85	84	94	87	91	88	92	89	88	96	95	89	86	94	90	92	92	94	89
KARENIN	97	96	92	95	86	87	89	95	92	90	92	91	93	89	95	93	93	86	93	90	93	90	94	89
LEONAN	83	93	88	95	83	84	87	93	86	85	89	94	88	89	89	94	89	84	90	87	88	92	87	89
NATHAN	92	91	90	92	84	80	83	84	80	87	83	91	89	88	84	84	83	85	86	90	88	90	93	89
NAYLA A	96	98	95	95	86	82	87	95	86	90	89	93	91	89	93	93	83	88	91	89	89	92	91	89
NIHAYA	96	96	92	95	87	85	89	95	90	88	88	92	91	90	92	93	83	85	90	90	90	92	94	89
RADELA	91	92	88	95	85	85	84	93	86	90	88	95	93	92	91	91	86	86	87	90	88	92	95	89
RADITYA	84	93	83	95	80	82	79	93	82	90	86	95	89	90	91	93	82	84	84	87	87	89	91	89
RIZKY YL	94	93	91	95	86	86	92	95	89	86	90	92	90	89	95	90	89	87	93	90	90	89	94	89
SANDI AI	84	85	88	95	81	81	77	92	81	83	83	94	90	87	84	84	87	87	83	89	86	88	85	80
SHEVIKA	96	96	91	95	85	85	91	95	85	90	85	91	90	90	92	89	84	86	94	90	90	90	94	89
SONIA V	90	92	86	95	86	81	84	94	87	85	90	93	93	88	91	91	93	87	87	89	92	92	91	89
TALITA N	92	90	92	96	89	84	89	95	93	90	90	91	93	92	97	92	91	84	91	87	95	92	91	89
TARISA C	87	93	94	95	82	87	94	85	90	84	94	94	86	89	90	91	82	83	93	91	89	92	88	89
TEDY KU	88	92	88	95	83	82	80	91	81	85	85	94	88	88	86	85	86	86	85	89	85	92	91	89

Gambar 3 Data Nilai Siswa

Pengujian indeks validitas digunakan untuk mencari jumlah k yang optimal. Indeks DB adalah indeks validitas yang menghitung nilai rata-rata setiap titik dalam kumpulan data. Gambar 4 menunjukkan nilai indeks DB dari jumlah *cluster* tiga sampai 10. Dari gambar tersebut terlihat bahwa nilai $k=3$ memiliki nilai indeks DB terkecil, yaitu sebesar 0.074. Semakin kecil nilai DBI yang dihasilkan (non-negatif), maka didapatkan jumlah *cluster* yang semakin baik.



Gambar 4 Nilai Indeks DB

Perhitungan *k-means clustering* dijelaskan sebagai berikut:

1. Menentukan jumlah *cluster*

Jumlah *cluster* yang akan digunakan pada data siswa sebanyak tiga *cluster* yaitu pintar, sedang, dan cukup berdasarkan nilai siswa kelas IXA - IXH. Setelah proses pengumpulan data siswa dilakukan maka selanjutnya menerapkan proses algoritma *k-means*. Algoritma *k-means* diterapkan pada data penelitian berjumlah 266 siswa dan 24 atribut, yang mencakup 12 nilai pengetahuan dan 12 nilai keterampilan dari 12 mata pelajaran. Parameter data dapat dilihat pada Tabel 1.

Tabel 1 Parameter Data

Jumlah <i>Cluster</i>	Definisi	Atribut	Jumlah siswa
3	Pintar, sedang, dan cukup	24	266

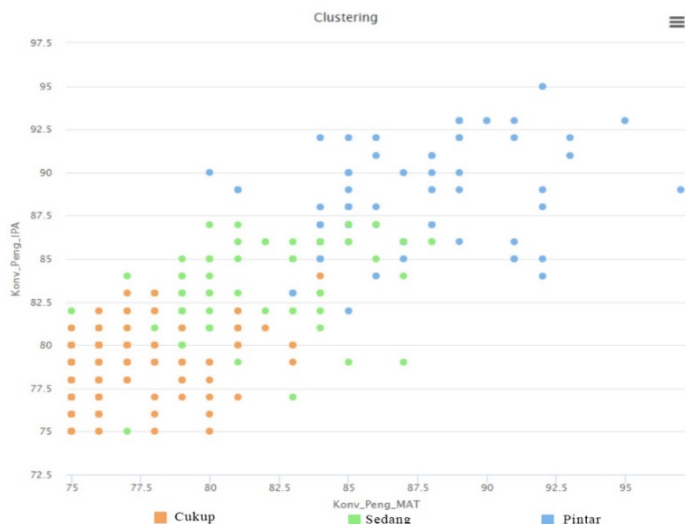
2. Menentukan *centroid*

Tabel 2 menampilkan nilai titik pusat setiap *cluster*. Setelah mendapatkan titik pusat setiap *cluster*, jarak tiap data dengan pusat *cluster*-nya dihitung ulang, kemudian dibentuk pusat *cluster* baru sampai *centroid cluster* yang diperoleh tidak bergerak. Dalam penelitian ini perhitungan dilakukan sampai tujuh iterasi, sampai tidak ditemukan adanya perubahan *centroid* untuk masing-masing *cluster*. Plot hasil merupakan pemetaan grafik dari hasil perhitungan *clustering*. Hasil ini dapat menentukan makna dari nilai k . Plot pada Gambar 5 merupakan *scatter plot* antara nilai pengetahuan IPA dan Matematika. Dari *scatter plot* tersebut dapat diketahui

bahwa *cluster* yang pintar adalah *cluster* yang berwarna biru, *cluster* sedang berwarna hijau, dan *cluster* yang cukup berwarna orange.

Tabel 2 Nilai *Centroid*

Atribut	Pintar	Sedang	Cukup
Peng_PABK	92.056	87.837	85.155
Ketr_PABK	93.648	90.359	86.738
Peng_PPKN	91.056	86.880	84.709
Ketr_PPKN	94.014	88.533	85.282
Peng_IND	84.789	82.815	81.087
Ketr_IND	85.493	85.891	83.709
Peng_MAT	87.070	80.076	77.068
Ketr_MAT	93.831	86.913	80.757
Peng_IPA	88.437	82.109	78.583
Ketr_IPA	86.521	83.500	81.718
Peng_IPS	87.549	85.783	84.660
Ketr_IPS	91.380	90.783	89.447
Peng_ING	89.606	86.109	79.699
Ketr_ING	88.324	86.326	80.883
Peng_SENI	92.592	86.880	83.049
Ketr_SENI	90.915	88.674	86.417
Peng_PENJAS	86.225	83.337	83.893
Ketr_PENJAS	85.690	85.011	86.155
Peng_PRAK	92.225	88.500	87.078
Ketr_PRAK	88.598	89.457	89.320
Peng_JAWA	92.028	87.435	83.573
Ketr_JAWA	90.535	87.283	84.524
Peng_TIK-BK	92.789	90.304	88.272
Ketr_TIK-BK	88.380	85.761	86.214



Gambar 5 Scatter Plot

Tabel 3 Anggota Masing-masing Cluster

Cluster pintar		Cluster sedang		Cluster cukup	
No	Nama	No	Nama	No	Nama
1	AJN	1	ANH	1	DS
2	ABN	2	AFI	2	DS
3	ASA	3	DJP	3	EDD
4	AKL	4	NAR	4	FPS
5	AHB	5	REK	5	FAA
6	AC	6	SAF	6	IPW
7	ADL	7	TK	7	JO
8	AAT	8	ANA	8	JV
9	APA	9	ARL	9	JP
10	BFF	10	IAW	10	LSJ
...
62	DVB	83	LGA	94	ILA
63	HFS	84	LDA	95	MRD
64	INM	85	NAR	96	MAP
65	LIA	86	PAW	97	MZA
66	NAL	87	RFK	98	PS
67	NKD	88	SAM	99	PVS
68	SA	89	SCA	100	PDE
69	SSR	90	SAD	101	RKW
70	WKD	91	VLK	102	RAR
71	YCN	92	WLS	103	TC

Berdasarkan hasil perhitungan *k-means* menggunakan *software* RapidMiner, didapatkan analisis yaitu didapatkannya tiga kelompok belajar yaitu pintar, sedang dan cukup. Jumlah kelompok belajar setiap kategori *cluster* menghasilkan hasil yang berbeda. Tabel 3 merupakan anggota masing-masing

cluster, dimana 71 siswa termasuk dalam *cluster* pintar, 92 siswa termasuk dalam *cluster* sedang, dan 103 siswa termasuk dalam *cluster* siswa cukup.

Kurikulum 2013 yang digunakan saat ini menerapkan penilaian terhadap dua kategori nilai, yaitu nilai pengetahuan dan keterampilan, sedangkan kurikulum sebelumnya hanya didasarkan pada nilai pengetahuan saja. Dengan sumber data yang sama, dilakukan perbandingan antara pengelompokan menggunakan nilai pengetahuan dan nilai keterampilan dan pengelompokan menggunakan nilai pengetahuan saja. Perbandingan ini menggunakan parameter yang sama, hanya saja untuk pengelompokan dengan nilai pengetahuan menggunakan 12 atribut dengan menghilangkan atribut nilai keterampilan.

Tabel 4 merupakan hasil *clustering* dengan nilai pengetahuan dan keterampilan, dimana *cluster* pintar terdiri dari 71 siswa, *cluster* sedang terdiri dari 92 siswa, dan *cluster* cukup terdiri dari 103 siswa. Sementara itu, Tabel 5 merupakan hasil *clustering* dengan nilai pengetahuan saja, dimana *cluster* pintar terdiri dari 72 siswa, *cluster* sedang terdiri dari 87 siswa, dan *cluster* cukup terdiri dari 107 siswa. Adanya pengurangan dan penambahan di setiap kategori *cluster* jika dilakukan perbandingan untuk perhitungan tanpa nilai keterampilan, *cluster* pintar terdapat penambahan sebanyak satu siswa, *cluster* sedang terdapat pengurangan siswa sebanyak lima siswa, dan *cluster* cukup terdapat penambahan siswa sebanyak empat siswa. Hasil perbandingan ini mendapatkan analisis bahwa nilai keterampilan berpengaruh terhadap pergerakan *cluster* pintar, sedang dan cukup. Walaupun pergerakan tidak terlalu signifikan tetapi tetap bisa dijadikan acuan terhadap pengelompokan kelompok belajar.

Tabel 4 Hasil *Clustering* Nilai Pengetahuan dan Keterampilan

<i>Cluster</i>	Jumlah siswa
<i>Cluster</i> pintar	71
<i>Cluster</i> sedang	92
<i>Cluster</i> cukup	103

Tabel 5 Hasil *Clustering* Nilai Pengetahuan Tanpa Keterampilan

<i>Cluster</i>	Jumlah siswa
<i>Cluster</i> pintar	72
<i>Cluster</i> sedang	87
<i>Cluster</i> cukup	107

Lebih lanjut pada Tabel 6 disajikan perubahan jumlah anggota tiap *cluster* kelompok belajar antara perhitungan *clustering* dengan nilai pengetahuan dan keterampilan dan perhitungan dengan nilai pengetahuan saja. Perpindahan siswa-siswa tersebut dikarenakan adanya perbedaan atribut yang digunakan, dimana pada penilaian pengetahuan dan keterampilan terdapat 24 atribut, sedangkan pada penilaian pengetahuan saja terdapat 12 atribut.

Tabel 6 Perbandingan Nilai Pengetahuan dan Keterampilan (PK) dengan Nilai Pengetahuan Tanpa Keterampilan (P)

PK \ P	P		
	Pintar	Sedang	Cukup
Pintar	66	5	0
Sedang	6	76	10
Cukup	0	6	97

Sebanyak 66 siswa tetap berada di *cluster* pintar, 76 siswa tetap berada di *cluster* sedang, dan 97 siswa tetap berada di *cluster* cukup, baik pada penilaian PK maupun P. Hal ini disebabkan nilai keterampilan dan nilai pengetahuannya tidak mempengaruhi penilaian untuk berpindah *cluster* atau bisa dikatakan nilai yang didapatkan di nilai PK dan P adalah sama.

Sebanyak lima siswa pada penilaian PK berada pada *cluster* pintar ternyata pada penilaian P termasuk *cluster* sedang dan sebanyak 10 siswa pada penilaian PK berada pada *cluster* sedang, tetapi pada penilaian P termasuk *cluster* cukup. Hal ini disebabkan karena selain pengetahuan mereka bagus, mereka juga mempunyai nilai lebih pada keterampilan, sehingga ketika dilakukan penilaian dengan PK, prestasi mereka juga semakin meningkat.

Sebanyak enam siswa pada penilaian PK berada pada *cluster* sedang, tetapi pada penilaian P berada pada *cluster* pintar dan sebanyak enam siswa pada penilaian PK berada pada *cluster* cukup, tetapi pada penilaian P berada pada *cluster* sedang. Hal ini disebabkan karena siswa-siswa tersebut mempunyai kemampuan yang baik di pengetahuan saja, sehingga ketika nilai keterampilan, yang sebenarnya merupakan penerapan dari nilai pengetahuan, dimasukkan sebagai poin penilaian, siswa-siswa tersebut berkurang penilaiannya.

Dari perbandingan ini dapat dilihat bahwa siswa meningkat prestasinya ketika nilai pengetahuan dan keterampilan dijadikan patokan penilaian akademik siswa, dibandingkan ketika nilai pengetahuan saja yang dijadikan patokan. Oleh karena itu, kurikulum 2013 sudah tepat untuk dijadikan standar pendidikan saat ini.

5. Simpulan

Penerapan algoritma *k-Means* membagi *dataset* penilaian siswa yang meliputi nilai pengetahuan dan keterampilan menjadi tiga kelompok berdasarkan indeks DB, yaitu *cluster* pintar, sedang, dan cukup. Berdasarkan perhitungan *clustering* dengan nilai pengetahuan dan keterampilan didapatkan kelompok pintar dengan jumlah 72 siswa, sedang dengan jumlah 92 siswa, dan cukup dengan jumlah 103 siswa. Sementara itu, perhitungan *clustering* dengan nilai pengetahuan saja didapatkan kelompok pintar dengan jumlah 71 siswa, sedang dengan jumlah 87 siswa, dan cukup dengan jumlah 107 siswa. Dari perbandingan pengelompokan dengan nilai pengetahuan dan keterampilan dan pengelompokan dengan nilai pengetahuan saja didapatkan hasil bahwa terjadi perpindahan *cluster* siswa yang

tidak terlalu signifikan, yaitu sebesar 10.15 persen dari keseluruhan siswa, dimana pengelompokan dengan nilai pengetahuan dan keterampilan lebih baik dibandingkan pengelompokan dengan nilai pengetahuan saja.

Hasil pengelompokan ini dapat digunakan pihak sekolah dalam membantu menentukan sistem pengajaran dengan memperhatikan nilai siswa melalui dua kategori nilai tersebut, sehingga siswa yang unggul di pengetahuan, tetapi tidak unggul di keterampilan mendapatkan pengajaran yang akan menyeimbangkan nilai akademik pengetahuan dan keterampilan. Penelitian ini dapat dikembangkan dengan menambahkan penelitian yang dilakukan dengan analisis prestasi siswa tidak hanya berdasarkan nilai akademis saja, tetapi juga menggunakan nilai non akademis seperti nilai ekstrakurikuler. Selain itu dapat dilakukan analisis berdasarkan nilai siswa menggunakan metode lain seperti *Fuzzy C-Means* dan melakukan perbandingan antara perhitungan *clustering* dengan metode *k-means* dan *Fuzzy C-Means*.

Daftar Pustaka

- [1] S. Budi, *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*, Ed. 1. Yogyakarta: Graha Ilmu, 2007.
- [2] S. Agustina, D. Yhudo, H. Santoso, N. Marnasusanto, A. Tirtana, and F. Khusnu, "Clustering Kualitas Beras Berdasarkan Ciri Fisik Menggunakan Metode K-Means Algoritma," *Clust. K-Means*, pp. 1–7, 2012.
- [3] G. Abdurrahman, "Clustering Data Ujian Tengah Semester (UTS) Data Mining Menggunakan Algoritma K-Means," *J. Sist. dan Teknol. Inf. Indones.*, vol. 1, no. 1, pp. 71–79, 2016.
- [4] A. I. Warnilah, "Analisis Algoritma K-Means Clustering Untuk Pemetaan Prestasi Siswa," *Indones. J. Comput. Inf. Technol.*, vol. 1, no. 1, pp. 83–95, 2016.
- [5] A. S. Kusuma and K. S. Aryati, "Sistem Informasi Akademik Serta Penentuan Kelas Unggulan Dengan Metode Clustering Dengan Algoritama K-Means Di Smp Negeri 3 Ubud," *J. Sist. Inf. dan Komput. Terap. Indones.*, vol. 1, no. 3, pp. 143–152, 2019.
- [6] C. Purnamaningsih, R. Saptono, and A. Aziz, "Pemanfaatan Metode K-Means Clustering dalam Penentuan Penjurusan Siswa SMA," *J. Teknol. Inf. ITS smart*, vol. 3, no. 1, pp. 27–33, 2016.
- [7] M. A. W. K. Murti, "Penerapan Metode K-Means Clustering Untuk Mengelompokan Potensi Produksi Buah – Buah Di Provinsi Daerah Istimewa Yogyakarta," Universitas Sanata Dharma, 2017.
- [8] Ediyanto, N. Mara, and N. S. Intisari, "Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis," *Bul. Ilm. Mat. Stat. dan Ter.*, vol. 02, no. 2, pp. 133–136, 2013.
- [9] Y. S. Thakare and S. B. Bagal, "Performance Evaluation of K-means Clustering Algorithm with Various Distance Metrics," *Int. J. Comput. Appl.*, vol. 110, pp. 12–16, 2015.
- [10] E Mulyasa, *Pengembangan dan Implementasi kurikulum 2013*. Bandung: Remaja Rosadakarya, 2017.

- [11] C. K. R. Charu C. Aggarwal, *Data Clustering: Algorithms and Applications*. Florida: Chapman and Hall/CRC, 2014.
- [12] J. P. Jiawei Han, Micheline Kamber, *Data Mining: Data Mining Concepts and Techniques*. Massachusetts: Morgan Kaufmann, 2011.