

Clustering Data Menggunakan Metode K-Means untuk Rekomendasikan Pembelajaran Akademik bagi Siswa Aktif dalam Ekstrakurikuler

Muhammad Azzam Al Fauzie, Yuliadi*, Juniardi Akhir Putra

Fakultas Rekayasa Sistem, Program Studi Informatika, Universitas Teknologi Sumbawa, Sumbawa, Indonesia

Email: ¹azzamliwa25@gmail.com, ^{2,*}yuliadi@uts.ac.id, ³juniardi.akhir.putra@uts.ac.id

Email Penulis Korespondensi: yuliadi@uts.ac.id

Abstrak—Prestasi akademik siswa meliputi pencapaian dan kinerja mereka dalam bidang akademik, seperti hasil ujian, nilai rapor, dan prestasi dalam pelajaran tertentu. Prestasi akademik siswa tersebut berpengaruh dalam studi lanjut ke tingkat yang lebih tinggi. Pada penelitian ini dilakukan pengelompokan data siswa menggunakan algoritma *K-Means Clustering*. Data yang digunakan adalah nilai rapor siswa kelas VIII dan IX. Hasil pengelompokan data menunjukkan adanya delapan *cluster* dengan jumlah siswa yang berbeda-beda dengan menggunakan tiga kategori ukuran *cluster* yaitu pintar, sedang, dan cukup. Berdasarkan analisis data, ditemukan bahwa Bahasa Inggris perlu mendapatkan perhatian khusus dalam pembelajaran. Siswa dengan nilai seimbang antara mata pelajaran IPA dan IPS dapat memilih jurusan sesuai minat mereka. Siswa dalam *cluster* sedang cenderung memiliki nilai yang dominan pada jurusan IPA, sementara siswa dalam *cluster* cukup cenderung memiliki nilai yang dominan pada jurusan IPS. Mayoritas siswa memiliki nilai akademik di atas 80, menunjukkan kinerja yang baik. Dari pengelompokan data tersebut menjadi dasar dalam pemilihan jurusan yang sesuai, partisipasi optimal dalam kegiatan ekstrakurikuler, dan dukungan bagi siswa dengan nilai cukup. Diharapkan hasil penelitian menjadi rekomendasi bagi sekolah untuk meningkatkan pemahaman siswa dan menjaga keseimbangan antara akademik dan ekstrakurikuler.

Kata Kunci: Metode Pembelajaran; Nilai Akademik; Ekstrakurikuler; K-Means Clustering; Data Mining; Rekomendasi Jurusan

Abstract—Student academic achievement includes their achievements and performance in academic fields, such as exam results, report card scores, and achievements in certain subjects. The student's academic achievement has an effect on further study to a higher level. In this study, grouping of student data was carried out using the *K-Means Clustering* algorithm. The data used are grades VIII and IX grade report cards. The results of grouping the data showed that there were eight clusters with different numbers of students using three categories of cluster sizes, namely smart, moderate, and moderate. Based on data analysis, it was found that English needs to get special attention in learning. Students with balanced grades between science and social studies can choose a major according to their interests. Students in the moderate cluster tend to have dominant grades in the Science major, while students in the moderate cluster tend to have dominant grades in the Social Sciences major. The majority of students have grades above 80, indicating good performance. The grouping of these data forms the basis for selecting appropriate majors, optimal participation in extracurricular activities, and support for students with sufficient grades. It is hoped that the research results will become a recommendation for schools to improve students' understanding and maintain a balance between academic and extracurricular activities.

Keywords: Learning Methods; Academic Performance; Extracurricular; K-Means Clustering; Data Mining; Major Recommendations

1. PENDAHULUAN

Ekstrakurikuler adalah latihan kurikuler yang dilakukan oleh siswa di luar konsentrasi jam, latihan intrakurikuler dan latihan kokurikuler, di bawah arahan dan pengelola satuan persekolahan [1][2]. Kegiatan ekstrakurikuler adalah merupakan kegiatan di luar jam pelajaran biasa yang bertujuan untuk memperluas pengetahuan, menyalurkan bakat dan minat siswa. Selain itu, ekstrakurikuler juga dapat membantu siswa dalam mempersiapkan diri untuk masa depan, seperti dengan belajar tentang bisnis atau profesi tertentu melalui ekstrakurikuler bisnis atau profesi [3][4]. Ekstrakurikuler dapat berfungsi sebagai cara untuk meningkatkan informasi siswa dan secara signifikan mempengaruhi prestasi belajar secara keseluruhan. Ekstrakurikuler adalah rangkaian kesempatan belajar di luar kelas yang dirancang untuk memperluas wawasan siswa, menumbuhkan kembangkan bakat dan minat, serta menanamkan rasa pengabdian pada masyarakat [5][6].

MTsN 1 Lampung Barat sekolah yang berdiri pada tahun 1967 di desa Gunung Sugih Liwa, Kecamatan Balik Bukit, Kabupaten Lampung Barat. Seiring perkembangan zaman MTsN 1 Lampung Barat memiliki ekstrakurikuler yang masih aktif di sekolah hingga sekarang diantaranya terdapat : Pramuka, PMR, UKS, Keagamaan, Tahfidz, Paduan Suara, Seni (Tari, Paduan Suara, Drama) dan Bidang Olahraga [7]. Seiring dengan semakin berkembangnya dunia pendidikan, banyak sekolah menengah pertama yang menyediakan Ekstrakurikuler untuk mengasah skill siswanya khususnya di MTsN, dengan banyaknya ekstrakurikuler yang aktif di sekolah membuat banyak siswa tertarik untuk mengikutinya dan fokus pada ekstrakurikuler yang diikuti tersebut membuat nilai akademik siswa kurang baik atau terpengaruh. Berdasarkan hasil wawancara awal dengan Bapak Desembri, M.Pd. sebagai wakil kepala Kurikulum di MTsN 1 Lampung Barat, sekolah tersebut merupakan salah satu sekolah yang menawarkan berbagai kegiatan ekstrakurikuler yang sangat diminati siswa. Hal ini yang membuat pihak sekolah khawatir akan nilai kelulusan siswanya yang tidak memenuhi standar nantinya.

Salah satu cara penyampaian materi pembelajaran adalah melalui metode pembelajaran. Dengan memanfaatkan strategi pembelajaran, pengalaman yang berkembang dapat lebih sederhana dan lebih cepat untuk mencapai pemahaman [8]. Setelah melakukan hasil wawancara terdapat metode yang dilakukan dalam melakukan pembelajaran di MTsN 1 Lampung Barat terdiri dari Ceramah, Diskusi Kelompok, Demonstrasi, Tanya Jawab, Karyawisata, Discovery, Bermain Peran, Tugas Proyek, Problem Solving dan Video/Alef.

Berdasarkan permasalahan tersebut dibutuhkan sebuah solusi dalam mencari metode pembelajaran yang bisa mengimbangi nilai akademik siswa dan ekstrakurikuler yang diikutinya dan mengetahui nilai rata-rata siswa yang aktif mengikuti ekstrakurikuler, dengan cara *cluster* nilai akademik siswa berdasarkan nilai rapor kelas VIII dan IX menggunakan algoritma K-Means pada Data Mining. Dari *cluster* yang telah dilakukan nantinya juga pihak sekolah bisa mengetahui solusi bagi siswanya agar nilai rata-rata memenuhi standar kelulusan.

2. METODOLOGI PENELITIAN

2.1 Kajian Pustaka

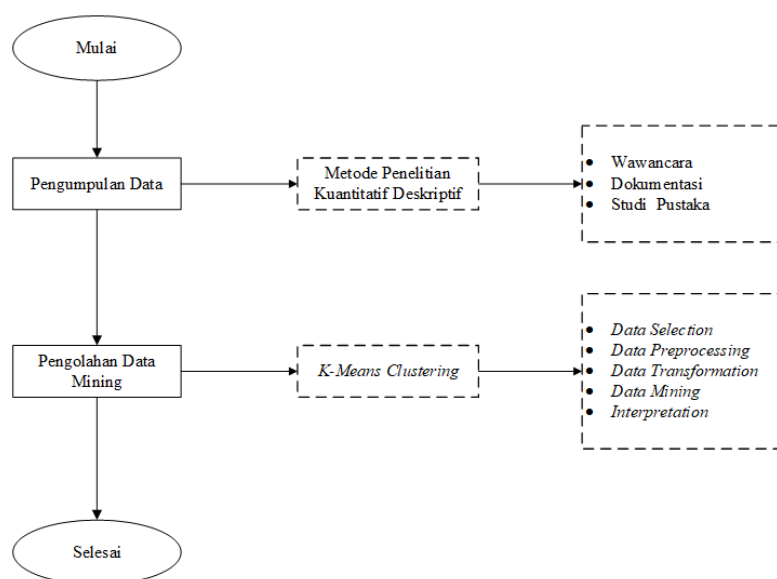
Beberapa penelitian sebelumnya yang mengulas tentang masalah yang berkaitan erat penggunaan metode *K-Means* adalah Penentuan Ekstra Kurikuler (Ekskul) Siswa SMK Terpadu Lampung Subang dengan Algoritma C 4.5 yang menghasilkan dari kumpulan 20 data, yang dibagi menjadi 70% data *training* dan 30% data *testing*, hasil evaluasi dan validasi dengan confusion matrix menunjukkan bahwa algoritma C4.5 memiliki tingkat akurasi sepintar 81,94% [9]. Selain itu algoritma *K-Means* di gunakan untuk menganalisis hasil belajar siswa yang dapat diimplementasikan terhadap data nilai raport [10]. Metode *K-Means* digunakan untuk memberikan solusi rekomendasi dalam penentuan jurusan di perguruan tinggi setelah lulus SMA [11]. Algoritma *K-Means* digunakan untuk merekomendasikan mata kuliah yang dipilih siswa sesuai dengan minat dan bakat siswa, sehingga siswa dapat belajar secara optimal [12].

Algoritma *K-Means* digunakan untuk mendapatkan informasi dari penggalian data mahasiswa sebagai rekomendasi bagi STIKES Perintis Padang dalam melakukan promosi [13]. Nilai raport VIII dan IX menggunakan perhitungan *K-Means* dengan pengelompokan informasi merupakan strategi yang digunakan untuk menemukan kualitas komparatif dalam suatu kumpulan sehingga dapat memberikan data yang bermanfaat. Algoritma pengelompokan data saat ini digunakan secara luas di berbagai bidang, termasuk bioinformatika, pemrosesan gambar, penambahan data, pengambilan keputusan, pengenalan pola, dan banyak lagi [14].

Data mining merupakan istilah yang merujuk pada proses penggalian informasi atau pengetahuan yang terdapat dalam basis data. Pertumbuhan ukuran basis data semakin pesat, bahkan terkadang menyulitkan sistem administrator untuk memahami informasi yang tersimpan atau mengaitkannya dengan pertanyaan yang muncul [15]. Knowledge Discovery in Database (KDD) adalah sebuah teknik untuk membentuk pola atau aturan dalam informasi. Informasi yang dihasilkan diperoleh dari data yang besar, juga dikenal sebagai tambang data dan menghasilkan data yang berpotensi bermanfaat. Iterasi dalam proses penambahan data ini disebut sebagai proses KDD [16].

2.2 Tahapan Penelitian

Pada Penulisan ini penulis menggunakan dua metode, yaitu metode pengumpulan data dan pengolahan Data Mining. Jenis penulisan ini menggunakan pendekatan kuantitatif deskriptif. Dalam penelitian ini, langkah-langkah pengumpulan data dan pengolahan data mining dapat di lihat pada gambar 1 dibawah ini:



Gambar 1. Alur Metode Penulisan

2.2.1 Metode Pengumpulan Data

Metode atau teknik pengumpulan data adalah metode pengumpulan data. Penulis menggunakan data sekunder dalam karya ini. Istilah "data sekunder" mengacu pada informasi yang dikumpulkan oleh orang lain selain penulis aslinya. Sebagian pintar waktu, catatan atau dokumen dari sumber tidak langsung seperti majalah, kutipan dan publikasi lainnya

merupakan data sekunder [17]. Beberapa tahapan dalam metode pengumpulan data yang dilakukan oleh peneliti, yakni antara wawancara, dokumentasi dan studi Pustaka.

2.2.2 Pengolahan Data Mining

Pengolahan Data Mining yang dilakukan penulisan ini yaitu, menggunakan tahapan dalam Knowledge Discovery in Database (KDD). Tujuan dari Knowledge Discovery in Database (KDD) adalah untuk menghasilkan informasi berdasarkan urutan yang telah ditentukan dan menemukan pola yang berguna dalam data [18][19][20]. KDD sendiri terdiri dari beberapa langkah yang berbeda dalam prosesnya yakni:

a. Data Selection

Penulis melakukan pemilihan data yang sesuai dari data nilai rapor berdasarkan atribut yang terdiri dari akidah-akhlak, al-quran-hadis, bahasa arab, fikih, bahasa inggris, bahasa indonesia, matematika, ilmu pengetahuan alam, ilmu pengetahuan sosial, pjok, ppkn, prakarya, seni budaya sejarah kebudayaan islam, bahasa lampung, total dan ekstrakurikuler.

b. Data Preprocessing

Penulis melakukan membersihkan data, dimana *field* yang dianggap tidak sesuai dengan kebutuhan penulisan dapat dihapus. Selain itu, format data diubah kembali untuk memastikan keseragaman format yang konsisten.

c. Data Transformation

Proses yang mengubah data agar dapat digunakan dan ditelusuri oleh tools *RapidMiner*. Dengan kata lain, transformasi data melibatkan pemetaan data yang rumit. Seperti mengubah data nama ekstrakurikuler menjadi bilangan angka sehingga mempermudah perhitungannya.

d. Data Mining

Penulis melakukan proses pengambilan pola dari data yang tersedia dengan bantuan tools *RapidMiner* menggunakan algoritma *K-Means Clustering*.

e. Interpretation/evaluation

Proses mengubah pola menjadi pengetahuan dengan tools *RapidMiner* sehingga dapat dimanfaatkan untuk mendukung pengambilan keputusan berupa pengelompokan data atau *cluster*.

3. HASIL DAN PEMBAHASAN

3.1 Data Selection

Data yang digunakan dalam melakukan penelitian ini yaitu gabungan data siswa kelas 8 dan 9 pada semeseter ganjil dan genap yang diperoleh berdasarkan dokumentasi dari Waka Kurikulum MTs Negeri 1 Lampung Barat. Jumlah seluruh siswa di MTs Negeri 1 Lampung Barat ada 447 siswa yang terdiri dari kelas 8 berjumlah 226 siswa dan kelas 9 berjumlah 221 siswa. Dikarenakan terdapat 447 siswa sehingga total keseluruhan data 894 dataset yang terdiri dari nilai akademik semester ganjil dan genap. Kemudian data diseleksi dan akan digunakan sebagai bahan rekomendasi metode pembelajaran yang mempengaruhi nilai akademik terhadap ekstrakurikuler.

Adapun atribut yang digunakan dalam penentuan metode pembelajaran yang mempengaruhi nilai akademik terhadap ekstrakurikuler adalah Akidah-Akhlak, Al-Quran-Hadis, Bahasa Arab, Bahasa Indonesia, Bahasa Inggris, Fikih, Ilmu Pengetahuan Alam, Ilmu Pengetahuan Sosial, Matematika, PJOK, PPKN, Prakarya, Seni Budaya, Sejarah Kebudayaan Islam, Bahasa Lampung, dan Ekstrakurikuler.

Tabel 1. Data Selection

No	Nama	QH	AA	FIK	SKI	PPKn	...	Ekstrakurikuler
1	Alfatha Radisya	89	85	85	89	84	...	PRAMUKA
2	Allegra Dini Nralika Putri	88	86	83	87	84	...	PRAMUKA
3	Aqila Alghifary	90	90	82	90	82	...	PRAMUKA
4	Atika Maharani	91	90	86	91	85	...	SENI
5	Atiya Reisy Ghadiza	90	90	82	90	82	...	PRAMUKA
6	Aulia Putri Ramadani	89	90	84	90	83	...	PRAMUKA
7	Bima Wahyu Rafael	90	88	83	90	84	...	PRAMUKA
8	Bintara Al Qausar	87	90	88	94	89	...	PRAMUKA
9	Chelsy Natasya Fitri	90	90	84	90	84	...	SENI
10	Chyka Allin Tia Aura	90	90	83	90	82	...	PMR
11	Dandi Anggara	92	91	85	92	86	...	PRAMUKA
12	Elsyahira Zalfa Putri	91	90	83	91	86	...	SENI
13	Fais Ihsan Rabbani	93	92	89	93	88	...	PRAMUKA
14	Fakhry Athar Radhitya	94	93	91	92	89	...	PRAMUKA
15	Ferdiansyah Rahmad Pikuloen	92	85	83	91	81	...	PRAMUKA

...
894	Ziyaad Fa`Iq Al Fakhri	86	86	83	82	86	...	PRAMUKA

3.2 Data Preprocessing

Dalam tahap *preprocessing*, dilakukan langkah-langkah untuk membersihkan data. Data yang telah dipilih divalidasi dan dikonfigurasi ulang agar formatnya tetap konsisten. Dalam proses ini, Data yang tidak memiliki nilai atau keterangan akan dibersihkan dari *dataset*.

Tabel 2. Data Preprocessing

No	Nama	QH	AA	FIK	SKI	PPKn	...	Ekstrakurikuler
1	Alfatha Radisya	89	85	85	89	84	...	PRAMUKA
2	Allegra Dini Nralika Putri	88	86	83	87	84	...	PRAMUKA
3	Aqila Alghifary	90	90	82	90	82	...	PRAMUKA
4	Atika Maharani	91	90	86	91	85	...	SENI
5	Atiya Reisy Ghadiza	90	90	82	90	82	...	PRAMUKA
6	Aulia Putri Ramadani	89	90	84	90	83	...	PRAMUKA
7	Bima Wahyu Rafael	90	88	83	90	84	...	PRAMUKA
8	Bintara Al Qausar	87	90	88	94	89	...	PRAMUKA
9	Chelsy Natasya Fitri	90	90	84	90	84	...	SENI
10	Chyka Allin Tia Aura	90	90	83	90	82	...	PMR
11	Dandi Anggara	92	91	85	92	86	...	PRAMUKA
12	Elsyahira Zalfa Putri	91	90	83	91	86	...	SENI
13	Fais Ihsan Rabbani	93	92	89	93	88	...	PRAMUKA
14	Fakhry Athar Radhitya	94	93	91	92	89	...	PRAMUKA
15	Ferdiansyah Rahmad Pikuloen	92	85	83	91	81	...	PRAMUKA
...
894	Ziyaad Fa`Iq Al Fakhri	86	86	83	82	86	...	PRAMUKA

3.3 Data Transformation

Ketika menggunakan metode *K-Means* clustering untuk mengolah data di atas, perlu dilakukan transformasi data terlebih dahulu. Khususnya untuk data yang memiliki jenis nominal seperti ekstrakurikuler, data tersebut perlu diinisialisasikan dalam bentuk angka.

Dalam proses ini, setiap kategori ekstrakurikuler akan diberikan kode atau angka yang mewakili masing-masing kategori. Misalnya, ekstrakurikuler "Pramuka" dapat diinisialisasikan dengan angka 1, ekstrakurikuler "PMR" dapat diinisialisasikan dengan angka 2, dan seterusnya seperti tabel 3 dibawah ini.

Tabel 3. Hasil Inisialisasi Ekstrakurikuler

Ekstrakurikuler	Ket
Pramuka	1
PMR	2
UKS	3
Keagamaan	4
Tahfidz	5
Paduan Suara	6
Seni	7
Bidang Olahraga	8

Dengan melakukan inisialisasi ini, data yang semula berjenis nominal akan berubah menjadi data berjenis numerik, yang dapat digunakan dalam algoritma *K-Means* clustering untuk analisis lebih lanjut.

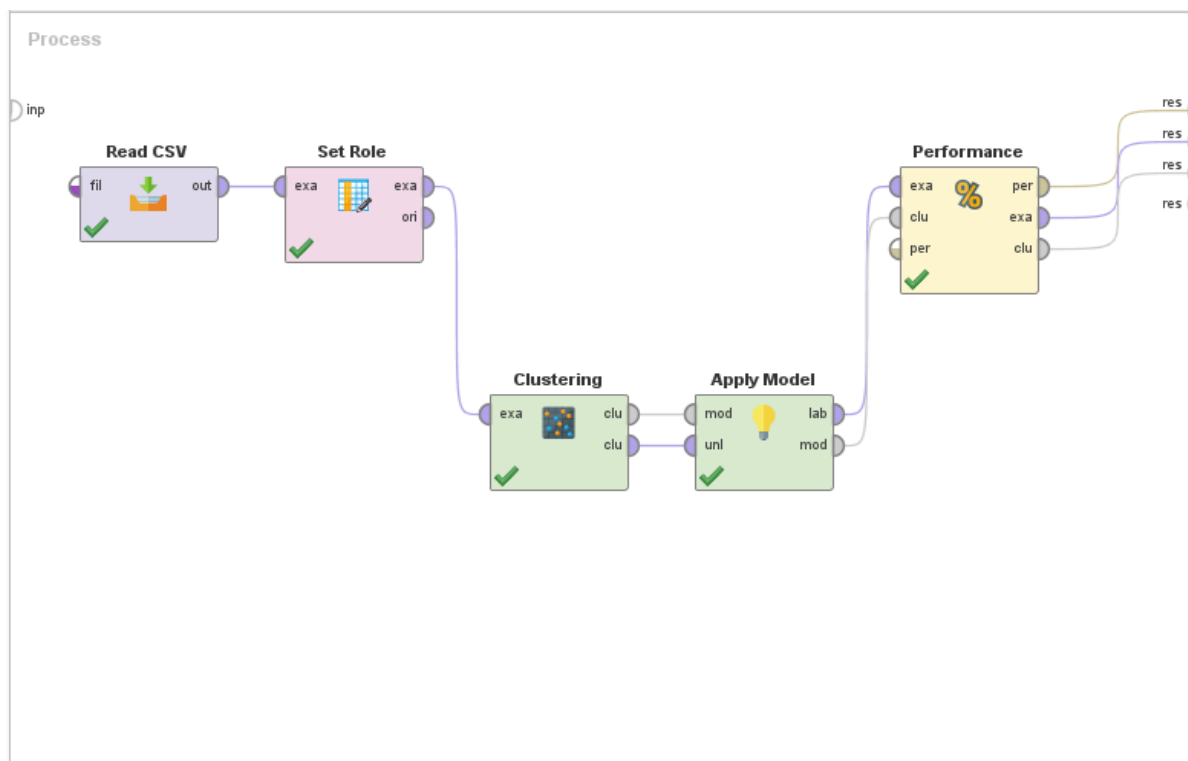
Tabel 3. Data Set

No	Nama	QH	AA	FIK	SKI	PPKn	...	Ekstrakurikuler
1	Alfatha Radisya	89	85	85	89	84	...	1
2	Allegra Dini Nralika Putri	88	86	83	87	84	...	1
3	Aqila Alghifary	90	90	82	90	82	...	1
4	Atika Maharani	91	90	86	91	85	...	7
5	Atiya Reisy Ghadiza	90	90	82	90	82	...	1
6	Aulia Putri Ramadani	89	90	84	90	83	...	1
7	Bima Wahyu Rafael	90	88	83	90	84	...	1
8	Bintara Al Qausar	87	90	88	94	89	...	1

9	Chelsy Natasya Fitri	90	90	84	90	84	...	7
10	Chyka Allin Tia Aura	90	90	83	90	82	...	2
11	Dandi Anggara	92	91	85	92	86	...	1
12	Elsyahira Zalfa Putri	91	90	83	91	86	...	7
13	Fais Ihsan Rabbani	93	92	89	93	88	...	1
14	Fakhry Athar Radhitya	94	93	91	92	89	...	1
15	Ferdiansyah Rahmad Pikuloen	92	85	83	91	81	...	1
...
894	Ziyaad Fa`Iq Al Fakhri	86	86	83	82	86	...	1

3.4 K-Means Clustering

Setelah mengumpulkan semua data siswa dari kelas 8 dan kelas 9, termasuk semester ganjil dan genap, data tersebut akan diubah menjadi bentuk angka. Selanjutnya, data-data ini dapat dikelompokkan menggunakan algoritma *K-Means Clustering*. Untuk melakukan proses pengolahan data ini, penulis menggunakan aplikasi *RapidMiner* untuk meminimalkan kesalahan yang mungkin terjadi selama proses pengolahan data. *RapidMiner* adalah solusi penambangan data, penambangan teks dan analisis prediktif yang tersedia sebagai perangkat lunak sumber terbuka. Untuk membantu pengguna membuat keputusan terbaik, perangkat lunak ini menggunakan berbagai metode deskriptif dan prediktif. Ada sekitar 500 operator penambangan data di *RapidMiner*, termasuk *operator input, output, preprocessing data*, dan visualisasi [12]. sehingga hasil data dari aplikasi dijamin keakuratan olahan data.

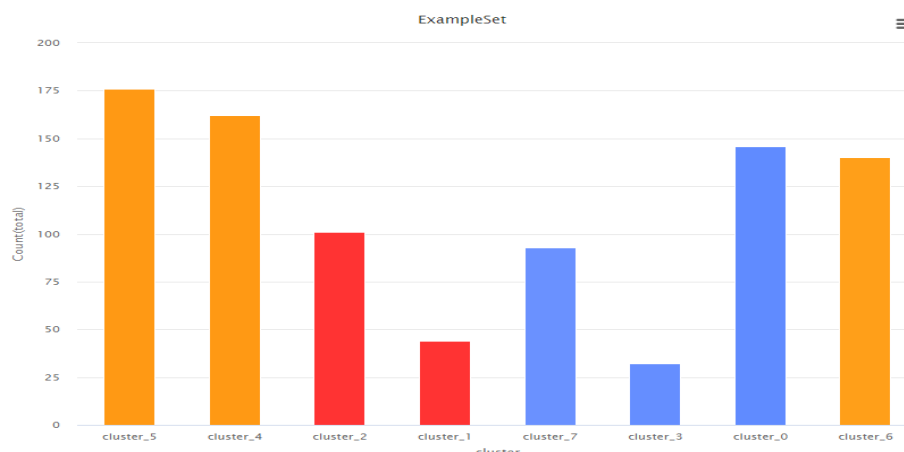


Gambar 2. Desain penggunaan algoritma *K-Means*

Gambar 2 Dalam penjelasan model CSV pertama, data CSV dibaca dan diolah menggunakan model. Model tersebut dapat diterapkan untuk membaca data CSV secara normal dan digunakan untuk prediksi. Semua parameter yang diperlukan disimpan dalam objek model. Dalam penelitian ini, penulis menggunakan sampel data siswa sebanyak 894 siswa dengan 17 atribut yaitu akidah-akhlak, al-quran-hadis, bahasa arab, bahasa indonesia, bahasa inggris, fikih, ilmu pengetahuan alam, ilmu pengetahuan sosial, matematika, pjok, ppkn, prakarya, seni budaya sejarah kebudayaan islam, bahasa lampung, total dan ekstrakurikuler.

3.5 Implementation Tools

Setelah melakukan input data dan konfigurasi *cluster* yang telah dilakukan, diperoleh hasil yang didapatkan 8 *cluster*, yakni *cluster 0* merupakan *cluster 1*, *cluster 1* merupakan *cluster 2*, *cluster 2* merupakan *cluster 3*, *cluster 3* merupakan *cluster 4*, *cluster 4* merupakan *cluster 5*, *cluster 5* merupakan *cluster 6*, *cluster 6* merupakan *cluster 7*, *cluster 7* merupakan *cluster 8* seperti pada gambar dibawah ini:



Gambar 4. Diagram Hasil Cluster

Hasil dari gambar 4 merupakan hasil dari clustering data siswa, yang di kelompokkan menjadi pintar (C2 dan C3), sedang (C5, C6 dan C7) dan cukup (C1, C4 dan C8) seperti warna yang ada pada gambar. Dalam hal ini nilai akademik siswa tecukup adalah Bahasa Inggris yang rata-rata nilai di bawah 80. Nilai pusat *cluster* terus berubah seiring dengan jumlah iterasi yang dilakukan. Perubahan ini terjadi karena jarak antara setiap data dengan pusat *cluster* yang diukur menggunakan jarak *euclidean* selalu berbeda pada setiap iterasi. Dalam konteks pengelompokan nilai akademik siswa, proses ini berakhir pada iterasi ke-8. Ini berarti bahwa nilai *cluster* pada iterasi ke-2, ke-4, ke-5, ke-6 dan ke-7 memiliki jumlah *cluster* yang sama, dengan nilai akademik bahasa Inggris memiliki nilai cukup di antara atribut lainnya. Berikut hasil grafik clusternya adalah:



Gambar 5. Grafik Hasil Clustering

Salah satu operator yang digunakan untuk mengukur kinerja *K-Means* adalah *accuracy %performance*. *Accuracy %performance* digunakan untuk mengevaluasi kinerja metode pengelompokan berbasis centroid. Parameter pengukuran kinerja termasuk *avg._within_centroid_distance* dan *davies bouldin*. *Avg._within_centroid_distance* mengukur rata-rata jarak antara data dengan pusat *cluster* dalam *cluster* yang sama, sedangkan *davies bouldin* adalah parameter yang mengukur kesamaan antara *cluster* dan ketidakterupaanannya. Algoritma pengelompokan dengan jarak antar *cluster* yang cukup dan kesamaan antar *cluster* yang cukup akan memiliki nilai *davies bouldin* yang cukup, yang menunjukkan performance yang lebih baik. Hasil pengukuran kinerja menggunakan indeks *davies bouldin* untuk pengelompokan nilai akademik siswa adalah 1.060. Nilai ini menunjukkan bahwa algoritma pengelompokan menghasilkan *cluster* yang memiliki jarak antar *cluster* yang cukup dan kesamaan antar *cluster* yang tinggi, sesuai dengan kriteria evaluasi. Oleh karena itu, berdasarkan hasil kinerja tersebut, dapat disimpulkan bahwa algoritma *K-Means* memiliki performance yang baik dalam pengelompokan nilai akademik siswa.

4. KESIMPULAN

Pengelompokan data siswa menggunakan *K-Means Clustering*, terdapat delapan *cluster* dengan jumlah data sebagai berikut: *Cluster 1* (146 data), *Cluster 2* (44 data), *Cluster 3* (101 data), *Cluster 4* (32 data), *Cluster 5* (162 data), *Cluster 6* (176 data), *Cluster 7* (140 data) dan *Cluster 8* (93 data). Setelah dilakukan perbersihan data dengan menghapus informasi nama siswa, jumlah siswa dalam setiap *cluster* berubah menjadi: di kelompokkan menjadi pintar (C2 (35 siswa)

dan C3 (70 siswa)), sedang (C5 (84 siswa), C6 (183 siswa) dan C7 (53 siswa)) dan cukup (C1 (48 siswa), C4 (25 siswa) dan C8(33 siswa)). Hasil proses analisis dilakukan dari data siswa yang aktif dalam kegiatan ekstrakurikuler di MTs Negeri 1 Lampung Barat, Terdapat beberapa *cluster* dengan nilai centroid Bahasa Inggris yang cukup. Diperlukan penerapan metode pembelajaran yang efektif, seperti simulasi, untuk meningkatkan pemahaman siswa dalam Bahasa Inggris. Penting bagi siswa-siswi dalam *cluster* cukup untuk mengoptimalkan keaktifan dalam kegiatan ekstrakurikuler sesuai minat mereka dan melibatkan diri dalam kegiatan bimbingan untuk meningkatkan pemahaman dalam pembelajaran. Secara umum, mayoritas siswa-siswi memiliki nilai akademik di atas 80, menunjukkan kinerja yang baik dalam pembelajaran. Hanya sejumlah kecil siswa yang memiliki rata-rata nilai di bawah 80. Oleh karena itu, tidak perlu khawatir tentang penurunan nilai akademik secara signifikan di kalangan siswa-siswi yang aktif dalam kegiatan ekstrakurikuler.

REFERENCES

- [1] Permendikbud, "Permendikbud Nomor 62 Tahun 2014," 2014, Accessed: Mar. 30, 2023. [Online]. Available: <https://jdih.kemdikbud.go.id/sjdih/siperpu/dokumen/salinan/Permendikbud%20Nomor%2062%20Tahun%202014.pdf>
- [2] A. W. Rifki and Listyaningsih, "Hubungan Kegiatan Ekstrakurikuler Pecinta Alam dengan Sikap Peduli Lingkungan Siswa di SMK Negeri 2 Bojonegoro," *Kaji. Moral dan Kewarganegaraan*, vol. 05, no. 20, pp. 426–440, 2017, [Online]. Available: <https://www.google.co.id/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiO99HY-quAAxXqXWwGHeD1CQU4ChAWegQIEhAB&url=https%3A%2F%2Fjournal.unesa.ac.id%2Findex.php%2Fjournal-pendidikan-kewarganegaraan%2Farticle%2Fdownload%2F18452%2F16826%2F>
- [3] R. Gupitha, "Penentuan Ekstra Kurikuler (EKSKUL) Siswa SMK Terpadu Lampung Subang Dengan Algoritma C 4.5," *In Global*, vol. 4, no. 2, pp. 31–37, 2018, Accessed: Mar. 30, 2023. [Online]. Available: <https://core.ac.uk/download/pdf/275909776.pdf>
- [4] Nisrinah, "Pengelolaan Layanan Ekstrakurikuler di SMAIT Al-Fityan School Kabupaten Goawa," *Universitas Negeri Makasar*, 2022. [Online]. Available: <http://eprints.unm.ac.id/23888/1/SKRIPSI NISRINAH 1743041010.pdf>
- [5] K. Shilviana and T. Hamami, "Pengembangan Kegiatan Kokurikuler dan Ekstrakurikuler," *Palapa*, vol. 8, no. 1, pp. 159–177, 2020, doi: 10.36088/palapa.v8i1.705.
- [6] N. Sugiarto, "Aktivitas ekstrakurikuler dalam menunjang pembelajaran pai di kelas viii SMP Islam Darussalam Palangka Raya," *Institut Agama Islam Negeri Palangkaraya*, 2020. [Online]. Available: <http://digilib.iain-palangkaraya.ac.id/3146/%0Ahttp://digilib.iain-palangkaraya.ac.id/3146/1/Skripsi Nyono Sugiarto - 130111791.pdf>
- [7] T. P. Madrasah, "RKJM-MTsN 1 Lampung Barat," 2022. <https://lampung.kemendikbud.go.id/files/lampung/file/MTsN1LampungBarat/ablr1459909546.pdf>
- [8] N. Mufidah, "Metode Pembelajaran Al-Ashwat," *al Mahāra: Jurnal Pendidikan Bahasa Arab*, vol. 4, no. 2, pp. 199–218, Dec. 2018, doi: 10.14421/ALMAHARA.2018.042-03.
- [9] M. Tamba, "Analisa Clustering Laporan Prestasi Belajar Siswa/Siswi Menggunakan K-Means (Studi Kasus : Sekolah SMA N 21 Medan)," *Pelita Informatika: Informasi dan Informatika*, vol. 7, no. 1, pp. 92–96, 2018, Accessed: Mar. 30, 2023. [Online]. Available: <http://ejournal.stmik-budidarma.ac.id/index.php/pelita/article/view/893>
- [10] P. A. Jusia, F. M. Irfan, and K. Kurniabudi, "Clustering data untuk rekomendasi penentuan jurusan perguruan tinggi menggunakan metode *K-Means*," *ikraith-informatika*, vol. 3, no. 3, pp. 75–84, 2019, Accessed: Apr. 12, 2023. [Online]. Available: <http://journals.upi-yai.ac.id/index.php/ikraith-informatika/article/download/596/445>
- [11] T. R. Rivanthio, M. Ramdhani, and A. Sahi, "Penerapan Teknik Clustering Data Mining untuk Memprediksi Kesesuaian Jurusan Siswa (Studi Kasus SMA PGRI 1 Subang)," *Faktor Exacta*, vol. 13, no. 2, p. 125, Aug. 2020, doi: 10.30998/faktorexacta.v13i2.6588.
- [12] E. Rahmah, E. Haerani, A. Nazir, S. Ramadhani, and T. Informatika UIN Sultan Syarif Kasim Riau Jl Soebrantas Km, "Penerapan Algoritma K-Medoids Clustering Untuk Menentukan Strategi Promosi Pada Data Mahasiswa (Studi Kasus : Stikes Perintis Padang)," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 3, 2022, Accessed: Mar. 30, 2023. [Online]. Available: <https://ojs.serambimekkah.ac.id/jnkti/article/view/4355>
- [13] P. A. Jusia, "Face Recognition Menggunakan Metode Algoritma Viola Jones Dalam Penerapan Computer Vision," *Jurnal PROCESSOR*, vol. 11, no. 1, pp. 663–675, 2016, Accessed: Apr. 12, 2023. [Online]. Available: <https://ejournal.unama.ac.id/index.php/processor/article/view/316>
- [14] Y. Indrasari, "Efisiensi Saluran Distribusi Pemasaran Kopi Rakyat Di Desa Gending Waluh Kecamatan Sempol (IJEN Bondowoso)," *Jurnal Manajemen Pemasaran*, vol. 14, no. 1, pp. 44–50, 2020, Accessed: Mar. 29, 2023. [Online]. Available: <https://doi.org/10.9744/pemasaran.14.1.44-50>
- [15] M. Z. Deny Jollyta, William Ramdhan, *Konsep Data Mining dan Penerapan*, Deepublish. Yogyakarta: Deepublish, 2020.
- [16] S. Yudarwati, "Clusterisasi Pola Penjualan Obat Menggunakan Metode K-Means Clustering (Studi Kasus Di Apotek Pelangi Sei. Jering Teluk Kuantan)," *J. Perencanaan, Sains, Teknol. dan Komput.*, vol. 3, no. 1, pp. 148–162, 2020, [Online]. Available: <https://media.neliti.com/media/publications/440973-none-26a0bc65.pdf>
- [17] Dunham and Margaret H, *Data Mining Introductory and Advanced Topics*. Pearson Education India, 2003.
- [18] F. L. Sibuea and A. Saptia, "Pemetaan Siswa Berprestasi Menggunakan Metode *K-Means* Clustering," vol. 1, pp. 85–92, 2017, Accessed: Mar. 30, 2023. [Online]. Available: <https://doi.org/10.33330/jurteks.v4i1.28>
- [19] I. K. Juni Arta, G. Indrawan, and G. R. Dantes, "Data Mining Rekomendasi Calon Mahasiswa Berprestasi Di Stmik Denpasar Menggunakan Metode Technique for Others Reference By Similarity To Ideal Solution," *JST (Jurnal Sains dan Teknol.*, vol. 5, no. 2, pp. 11–21, 2017, doi: 10.23887/jstundiksha.v5i2.8549.
- [20] N. S. Luki Ardiantoro, Soffa Zahara, "Pemanfaatan Knowledge Data Discovery(KDD) Pada Pola Permainan Atlet Bulutangkis," *Explor. IT J. Keilmuan dan Apl. Tek. Inform.*, vol. 11, no. 1, pp. 1–6, 2019, doi: 10.35891/explorit.v11i1.1467