



Analisis Algoritma K-Means Clustering Dalam Pengelompokan Prestasi Belajar Siswa Menengah Atas (SMA)

Rahmah Dila^{1*}, Sarjon Defit², Syafri Arlis²

¹ Fakultas Ilmu Komputer, Magister Teknik Informatika, Universitas Putra Indonesia YPTK Padang, Padang, Indonesia

² Fakultas Ilmu Komputer, Doktor Teknologi Informasi, Universitas Putra Indonesia YPTK Padang, Padang, Indonesia

Email: ¹rahmahdila66@gmail.com, ²sarjon_defit@upiyptk.ac.id, ³syafri_arlis@upiyptk.ac.id

Email Penulis Korespondensi: rahmahdila66@gmail.com

Abstrak—Penggunaan media sosial yang meningkat dikalangan siswa SMA memberikan dampak positif maupun negatif terhadap prestasi akademik. Hal itu dapat dilihat dari perubahan pola belajar, tingkat konsentrasi, serta motivasi siswa dalam mengikuti kegiatan pembelajaran. Penelitian ini bertujuan untuk mengelompokkan prestasi belajar siswa berdasarkan tingkat penggunaan media sosial menggunakan algoritma K-Means Clustering. K-Means Clustering merupakan salah satu metode utama dalam data mining, yang merupakan Teknik pengelompokan data berdasarkan kesamaan karakteristiknya. Adapun parameter yang digunakan dalam menganalisis penelitian ini yaitu Durasi Media sosial (X1), waktu aktif (X2), platform utama (X3), tujuan utama (X4), waktu akses media sosial saat belajar (X5), kecanduan media sosial (X6), tingkat kecanduan media sosial (X7), jumlah grup belajar (X8) dan rata-rata akademik (X9). Berdasarkan metode K-Means Clustering telah mampu mengelompokkan siswa berdasarkan tingkat penggunaan media sosial. Hasil tersebut dapat terlihat dari kategori cluster C0 (Tinggi) beranggotakan 46 siswa menunjukkan siswa dengan penggunaan media sosial tinggi yang cenderung prestasi akademiknya rendah, C1 (sedang) beranggotakan 80 siswa menunjukkan siswa dengan penggunaan media sosial sedang yang cenderung prestasi akademiknya sedang, dan C2 (Rendah) beranggotakan 72 Siswa menunjukkan siswa dengan penggunaan media sosial rendah yang cenderung prestasi akademiknya tinggi. Kontribusi penelitian ini memberikan manfaat bagi siswa dengan membantu mereka memahami hubungan antara kebiasaan penggunaan media sosial dan prestasi belajar, sehingga dapat mendorong pengelolaan waktu yang lebih efektif.

Kata Kunci: Data Mining; K-Means Clustering; Penggunaan Media Sosial; Prestasi Belajar; Siswa Sekolah Menengah Atas

Abstract—The increased use of social media among high school students has a positive and negative impact on academic achievement. This can be seen from changes in learning patterns, concentration levels, and students' motivation in participating in learning activities. This study aims to classify student learning achievement based on the level of social media use using the K-Means Clustering algorithm. K-Means Clustering is one of the main methods in data mining, which is a technique of grouping data based on the similarity of its characteristics. The parameters used in analyzing this study are Social Media Duration (X1), Active Time (X2), Main Platform (X3), Main Goal (X4), Social Media Access Time While Learning (X5), Social Media Addiction (X6), Social Media Addiction Level (X7), Number of Study Groups (X8) and Academic Average (X9). Based on the K-Means Clustering method, it has been proven to be able to group students based on the level of social media use. These results can be seen from the cluster category C0 (High) with 46 students, C1 (medium) with 80 students, and C2 (Low) with 72 students. The contribution of this research benefits students by helping them understand the relationship between social media usage habits and learning achievement, so as to encourage more effective time management.

Keywords: Data Mining; K-Means Clustering; Social Media Use; Learning Achievement; Senior High School Students

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah membawa perubahan signifikan dalam kehidupan manusia, termasuk dalam dunia Pendidikan [1]. Salah satu fenomena yang muncul akibat kemajuan teknologi adalah meningkatnya penggunaan media sosial di kalangan remaja, khususnya siswa Sekolah Menengah atas (SMA), seperti Instagram, Tiktok, WhatsApp, dan Facebook, tidak hanya digunakan sebagai sarana komunikasi, tetapi juga sebagai media hiburan dan sumber informasi [2]. Namun, penggunaan media sosial yang tidak terkontrol berpotensi mempengaruhi prestasi belajar siswa [3].

Penggunaan media sosial dapat memberikan dampak positif maupun negatif terhadap prestasi akademik, tergantung pada intensitas, tujuan, serta cara siswa memanfaatkannya dalam mendukung kegiatan belajar maupun sekadar hiburan. [4]. Di satu sisi, media sosial dapat menjadi sarana berbagi ilmu, memperluas wawasan, serta meningkatkan keterampilan digital siswa [5]. Namun, di sisi lain, penggunaan media sosial yang berlebihan dapat menyebabkan gangguan konsentrasi, menurunkan efektivitas belajar, serta mengurangi waktu yang seharusnya digunakan untuk kegiatan akademik, sehingga diperlukan analisis untuk mengidentifikasi pola penggunaannya di kalangan siswa dan dampaknya terhadap prestasi belajar mereka [6].

Penelitian terdahulu yang berhubungan dengan penelitian ini menghasilkan 4 klaster, Dimana mengelompokkan 430 responden remaja berusia 14 hingga 23 tahun [7]. Selain itu, penelitian terdahulu untuk menganalisis aktivitas dan respond penggunaan media sosial yang menghasilkan 3 klaster [8]. Terakhir, penelitian terdahulu tentang klasifikasi video pada media sosial youtube menghasilkan 3 klaster dengan nilai akurasi sebesar 99% [9].

Implementasi dalam penentuan prestasi belajar menggunakan konsep clustering [10]. Clustering dirancang untuk mengelompokkan data atau objek ke dalam beberapa kelompok (cluster) berdasarkan kemiripan karakteristiknya, sehingga objek dalam satu cluster memiliki tingkat kesamaan yang tinggi, sedangkan objek antar cluster memiliki tingkat perbedaan yang signifikan [11]. Dalam konteks prestasi belajar, clustering memungkinkan pengelompokan siswa berdasarkan indikator tertentu, seperti durasi penggunaan media sosial, waktu belajar, motivasi, serta nilai akademik [12].



Metode K-Means Clustering dapat melakukan pengelompokan penentuan prestasi belajar siswa Sekolah menengah atas (SMA) berdasarkan tingkat penggunaan media sosial [13]. Algoritma ini bekerja dengan menentukan sejumlah kluster (K) yang diinginkan, kemudian menempatkan pusat kluster (centroid) secara acak pada tahap awal [14]. Selanjutnya, setiap data dihitung jaraknya terhadap setiap centroid menggunakan rumus Euclidean Distance, kemudian ditempatkan pada kluster dengan jarak terdekat [15]. Setelah semua data dikelompokkan, centroid setiap kluster diperbarui berdasarkan rata-rata nilai anggota kluster tersebut. Proses ini diulang hingga posisi centroid tidak lagi berubah atau perubahan yang terjadi sangat kecil [16].

Riwayat penelitian terdahulu yang telah membahas hubungan antara media sosial, perilaku remaja, dan prestasi belajar yang pertama penelitian yang dilakukan [17] menunjukkan bahwa intensitas penggunaan media sosial tinggi dapat memengaruhi prestasi akademik siswa SMA, terutama ketika waktu belajar tergantikan oleh aktivitas di media sosial. Selanjutnya, penelitian lainnya [18] menemukan bahwa perilaku remaja dalam menggunakan media sosial dapat dikategorikan menjadi tiga kelompok, yaitu penggunaan untuk hiburan, komunikasi, dan pembelajaran. Hasil penelitian ini menegaskan bahwa siswa yang sering menggunakan media sosial untuk tujuan edukatif cenderung memiliki prestasi yang lebih baik dibandingkan dengan mereka yang menggunakannya hanya untuk hiburan [19].

Selain itu, [20] meneliti pengaruh penggunaan media sosial terhadap motivasi belajar siswa SMA, Hasilnya menunjukkan adanya dua dampak yang berlawanan dimana media sosial dapat meningkatkan motivasi ketika digunakan sebagai sumber referensi belajar, namun dapat menurunkan konsentrasi dan kedisiplinan jika penggunaannya berlebihan. Penelitian terbaru oleh [21] juga membuktikan bahwa terdapat hubungan signifikan antara pola penggunaan media sosial dengan capaian akademik siswa. Dengan menggunakan metode analisis clustering, penelitian tersebut mengelompokkan siswa berdasarkan durasi penggunaan media sosial dan tingkat pencapaian nilai rata-rata akademik.

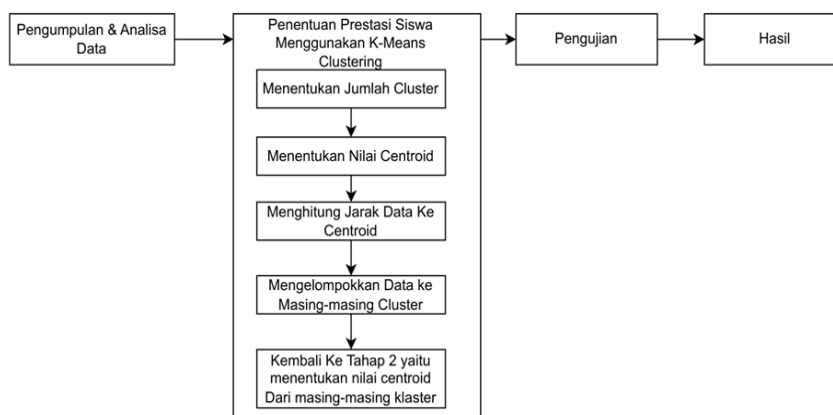
Dari penelitian-penelitian tersebut, dapat disimpulkan bahwa media sosial berpengaruh terhadap prestasi belajar siswa, baik secara positif maupun negatif. Namun, masih terbatas penelitian yang khusus menggunakan metode K-Means Clustering untuk mengelompokkan pola penggunaan media sosial siswa SMA dan mengaitkannya langsung dengan prestasi akademik. Hal ini menjadi pembeda utama penelitian ini dibandingkan dengan penelitian sebelumnya.

Berdasarkan penjelasan sebelumnya, maka penelitian ini akan melakukan penentuan prestasi belajar siswa menengah atas (SMA). Proses analisis tersebut dilakukan dengan menggunakan metode K-Means Clustering. K-Means Clustering dapat mengelompokkan data berdasarkan kesamaan karakteristik data. Hasil yang diharapkan dari penelitian ini adalah terbentuknya kluster yang merepresentasikan tingkat prestasi belajar siswa berdasarkan pola penggunaan media sosial. Penelitian ini berkontribusi dalam membantu siswa memahami keterkaitan antara pola penggunaan media sosial dengan prestasi belajar, sehingga dapat mendorong mereka untuk mengatur waktu secara lebih efisien. Selain itu, hasil penelitian juga bermanfaat bagi guru untuk memahami pola belajar siswa dan menyesuaikan strategi pembelajaran, bagi sekolah dalam merancang kebijakan literasi digital yang mendukung prestasi akademi, serta bagi orang tua dalam mengawasi penggunaan media sosial anak secara lebih bijak.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Proses penentuan prestasi belajar siswa menengah atas (SMA) dengan menggunakan algoritma K-Means Clustering, dibutuhkan sebuah kerangka kerja penelitian yang jelas dan sistematis. Kerangka kerja ini berfungsi sebagai acuan dalam menggambarkan alur penelitian secara teratah. Adapun kerangka penelitian ini dapat dilihat pada Gambar 1, berikut:



Gambar 1. Kerangka Kerja Penelitian

Pada Gambar 1 dapat dijelaskan Kerangka penelitian yang menjadi alur tahapan analisis penentuan prestasi siswa dengan metode K-Means Clustering. Proses dimulai dari pengumpulan dan analisis data, kemudian dilanjutkan dengan penentuan jumlah cluster, perhitungan centroid, hingga pengelompokan data ke dalam masing-masing cluster. Hasil dari proses ini selanjutnya diuji untuk memperoleh pola yang jelas mengenai hubungan intensitas penggunaan media sosial dengan capaian prestasi akademik siswa.



2.1.1 Pengumpulan dan Analisis Data

Pada Pengumpulan dan analisis data dimana data yang berisi parameter penggunaan media sosial dan nilai akademik siswa dikumpulkan dan dianalisis untuk memastikan kelengkapan serta kesesuaian dengan kebutuhan penelitian. Adapun parameter yang digunakan Durasi Media sosial (X1), waktu aktif (X2), platform utama (X3), tujuan utama (X4), waktu akses media sosial saat belajar (X5), kecanduan media sosial (X6), tingkat kecanduan media sosial (X7), jumlah grup belajar (X8) dan rata-rata akademik (X9). Variabel evaluasi (X9) yaitu nilai rata-rata akademik siswa digunakan sebagai salah satu variabel input dalam proses clustering, variabel ini juga berfungsi sebagai indikator evaluasi. Artinya, setelah kluster terbentuk, perbedaan nilai akademik antar kluster dianalisis lebih lanjut untuk melihat bagaimana hubungan pola penggunaan media sosial dengan prestasi belajar siswa. Proses pengumpulan data dilakukan dengan cara pemberian kuisioner kepada 198 siswa di SMA Negeri 7 Sijunjung.

2.1.2 K-Means Clustering

Algoritma K-Means merupakan salah satu teknik analisis clustering yang digunakan untuk membagi data ke dalam beberapa kelompok (cluster) berdasarkan kesamaan atribut, k-means dapat diadopsi di berbagai aplikasi dalam pengelompokan data sesuai karakteristik tertentu [22]. Adapun Langkah-langkah metode K-Means clustering sebagai berikut [23]:

a. Menentukan jumlah cluster (k).

b. Menentukan titik nilai centroid secara acak.

Menentukan nilai centroid awal dilakukan secara acak. Sedangkan menentukan nilai centroid iterasi berikut digunakan rumus:

$$C1 = (R1 + R2 + R3 + \dots + Rn) / (\sum_{i=1}^n [R]) \quad (1)$$

c. Menghitung jarak antara titik centroid dengan titik tiap objek untuk menghitung jarak tersebut dapat menggunakan Euclidean Distance, yaitu:

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

d. Pengelompokan data

Penentuan anggota cluster dilakukan dengan memperhitungkan jarak minimum objek. Nilai yang diperoleh dalam keanggotaan data pada distance matriks adalah 0 atau 1 untuk data yang dialokasikan ke cluster dan nilai 0 untuk data yang dialokasikan ke cluster yang lain.

e. Kembali ke tahap 2 yaitu menghitung Kembali nilai centroid baru dari masing-masing kluster, lakukan perulangan hingga nilai centroid yang dihasilkan tetap dan anggota cluster tidak berpindah ke cluster lain.

2.2 Pengujian

Hasil pengelompokan dianalisis untuk melihat ketepatan dan kesesuaian pola kluster dengan karakteristik data siswa. Pada bagian pengujian, proses dilakukan menggunakan Software analisis RapidMiner dalam memvisualisasikan proses kluster penentuan prestasi belajar siswa berdasarkan tingkat penggunaan media sosial di SMA Negeri 7 Sijunjung. Penggunaan software RapidMiner dapat mempermudah proses analisis dengan menyediakan antarmuka grafis yang intuitif dan mendukung interpretasi hasil secara lebih jelas dan akurat. Penentuan jumlah kluster optimal tidak hanya ditetapkan secara langsung, tetapi ditentukan melalui pendekatan metode Elbow yang dijalankan pada RapidMiner, sehingga diperoleh jumlah kluster yang paling representatif terhadap data. Selanjutnya, dilakukan validasi hasil clustering menggunakan metrik internal seperti Davies-Bouldin Index untuk memastikan bahwa hasil kluster stabil, terpisah dengan baik, dan tidak tumpang tindih.

Sebagai indikator evaluasi, penelitian ini tidak hanya melihat jumlah anggota kluster, tetapi juga menganalisis centroid tiap kluster, distribusi nilai akademik, serta profil pola penggunaan media sosial siswa pada setiap kelompok. Dengan demikian, hasil kluster dapat diinterpretasikan lebih lanjut, misalnya membedakan siswa dengan penggunaan media sosial berlebihan namun prestasi rendah, dibandingkan dengan siswa yang mampu mengatur penggunaan media sosial secara seimbang sehingga prestasi akademiknya lebih baik

3. HASIL DAN PEMBAHASAN

Pembahasan dan hasil penelitian ini disajikan untuk memberikan gambaran mengenai temuan yang diperoleh dari proses analisis data menggunakan algoritma K-Means Clustering. Bagian ini bertujuan untuk menjelaskan bagaimana data yang telah dikumpulkan diolah, dikelompokkan, serta ditafsirkan sehingga menghasilkan pola tertentu yang berkaitan dengan prestasi belajar siswa. Dengan demikian, pembahasan ini diharapkan mampu memberikan pemahaman yang lebih mendalam terhadap hubungan antara penggunaan media sosial dan capaian akademik siswa.

Hasil penerapan algoritma K-Means Clustering pada penelitian ini menghasilkan tiga kluster yang merepresentasikan tingkat penggunaan media sosial siswa SMA, yaitu kluster C0 (Penggunaan Media Sosial Tinggi, memiliki prestasi rendah) yang beranggotakan 46 data sampel siswa, C1 (Penggunaan Media Sosial Sedang, Memiliki Prestasi Sedang) yang beranggotakan 80 sampel data siswa, dan C2 (Penggunaan Media Sosial Rendah, memiliki prestasi



Tinggi) 72 sampel data. Adapun hasil k-means cluster dalam penentuan prestasi belajar siswa dapat dilihat pada Gambar 2.

Cluster Model

```
Cluster 0: 46 items
Cluster 1: 80 items
Cluster 2: 72 items
Total number of items: 198
```

Gambar 2. Hasil Cluster Model

3.1 Pengumpulan dan Analisis Data

Proses pengumpulan data dan analisis ditujukan untuk memastikan data yang dianalisis sesuai untuk menghasilkan hasil prestasi belajar siswa. Tahap ini merupakan bentuk gambaran awal mengenai variabel yang digunakan dalam penelitian ini. Penyajian sampel data ini bertujuan untuk memperlihatkan struktur serta karakteristik data yang menjadi dasar proses pengelompokan menggunakan algoritma K-Means Clustering, sehingga memudahkan pembaca dalam memahami konteks dan kualitas data sebelum dilakukan analisis lebih lanjut. Adapun dataset penelitian yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Dataset Penelitian

Nama	X1	X2	X3	X4	X5	X6	X7	X8	X9
Siswa 1	3	2	1	1	0	0	2	3	90
Siswa 2	2	2	2	2	0	0	1	2	88
Siswa 3	4	3	4	2	1	0	2	1	85
Siswa 4	6	4	3	2	1	1	4	2	78
Siswa 5	7	2	2	3	1	1	3	1	80
.....
Siswa 194	8	4	3	4	1	1	5	3	73
Siswa 195	2	1	1	1	0	0	1	2	94
Siswa 196	7	2	2	2	1	1	4	2	80
Siswa 197	8	4	3	4	1	1	5	3	73
Siswa 198	3	1	1	1	0	0	2	1	93

Tabel 1 menyajikan dataset penelitian yang digunakan dalam proses pengelompokan prestasi belajar siswa dengan algoritma K-Means Clustering. Dataset ini terdiri dari beberapa variabel yang direpresentasikan dalam sembilan indikator (X1–X9) yang menggambarkan karakteristik penggunaan media sosial serta rata-rata nilai akademik siswa. Kolom nama menunjukkan identitas responden, sedangkan setiap kolom indikator berisi nilai numerik hasil pengukuran terhadap kriteria penelitian, seperti durasi media sosial (X1), waktu aktif (X2), platform utama (X3), tujuan utama (X4), waktu akses media sosial saat belajar (X5), kecanduan media sosial (X6), tingkat kecanduan media sosial (X7), jumlah grup belajar (X8), dan rata-rata akademik (X9). Data ini menjadi dasar dalam proses analisis untuk melihat pola dan kecenderungan siswa dalam menggunakan media sosial serta hubungannya dengan prestasi belajar. Dengan demikian, tabel ini memberikan gambaran awal mengenai struktur data yang dianalisis sebelum dilakukan proses clustering lebih lanjut.

3.2 Hasil K-Means Clustering

Hasil penerapan algoritma K-Means Clustering pada penelitian ini menghasilkan tiga kluster yang merepresentasikan tingkat penggunaan media sosial siswa SMA, yaitu kluster C0 (Penggunaan Media Sosial Tinggi, memiliki prestasi rendah) yang beranggotakan 46 data sampel, C1 (Penggunaan Media Sosial Sedang, Memiliki Prestasi Sedang) yang beranggotakan 80 sampel data, dan C2 (Penggunaan Media Sosial Rendah, memiliki prestasi Tinggi) 72 sampel data. Adapun hasil k-means cluster dalam penentuan prestasi belajar siswa dapat dilihat pada Tabel 2.

Tabel 2. Hasil K-Means Cluster Dalam Penentuan Prestasi Belajar Siswa

Nama	C0	C1	C2	Min	Cluster
Siswa 3	5,258668	12,77438	7,75377	5,258668	0
Siswa 4	3,691693	5,280193	14,82524	3,691693	0
Siswa 5	2,062424	7,319569	12,86722	2,062424	0
.....
Siswa 8	8,903853	0,814275	20,35201	0,814275	1
Siswa 14	6,966247	2,776062	18,301	2,776062	1
Siswa 15	10,02265	2,228765	21,54347	2,228765	1
.....



Nama	C0	C1	C2	Min	Cluster
Siswa 1	9,871859	17,90707	2,49084	2,49084	2
Siswa 2	8,63589	16,41348	4,227405	4,227405	2
Siswa 6	10,38044	18,46295	2,926669	2,926669	2
.....

Tabel 2 menampilkan hasil pengelompokan data siswa menggunakan algoritma K-Means Clustering dalam menentukan prestasi belajar siswa berdasarkan tingkat penggunaan media sosial. Pada table terlihat bahwa setiap siswa memiliki nilai jarak terhadap masing-masing pusat klaster (C0,C1,C2). Kolom min menunjukkan nilai jarak terkecil yang menentukan klaster akhir siswa, sedangkan kolom cluster menunjukkan kelompok akhir tempat siswa tersebut ditempatkan. Dari hasil clustering, terbentuk tiga kelompok yaitu klaster C0 (Penggunaan Media Sosial Tinggi, Memiliki Prestasi Rendah) dengan jumlah 46 siswa, C1 (Penggunaan Media Sosial Sedang, Memiliki Prestasi Sedang) dengan 80 siswa, dan C2 (Penggunaan Media Sosial Rendah, memiliki prestasi Tinggi) dengan 72 siswa. Hal ini mengindikasikan bahwa terdapat perbedaan pola penggunaan media sosial yang berpengaruh terhadap pengelompokan prestasi belajar siswa, di mana siswa dengan jarak terkecil pada klaster tertentu otomatis tergolong kelompok tersebut. Agar lebih jelas dapat dilihat hasil ringkasan klasterisasi prestasi belajar pada Tabel 3.

Tabel 3. Ringkasan Hasil Klasterisasi Prestasi Belajar

Cluster	Jumlah Anggota	Ciri Utama	Rata-rata Akademik
Cluster 0	46 Siswa	Durasi Tinggi, sering akses saat belajar	72-77
Cluster 1	80 Siswa	Durasi sedang, penggunaan lebih seimbang	78-87
Cluster 2	72 Siswa	Durasi rendah, penggunaan terbatas, focus belajar	85-95

Berdasarkan Tabel 3, Cluster 0 didominasi oleh siswa dengan intensitas penggunaan media sosial yang tinggi, termasuk kebiasaan mengakses saat belajar, dengan rata-rata prestasi akademik yang relatif rendah (72-77). Cluster 1 merepresentasikan siswa dengan pola penggunaan media sosial sedang dan lebih terkontrol, dengan prestasi akademik menengah (78-87), sedangkan cluster 2 berisi siswa dengan intensitas penggunaan media sosial rendah, cenderung fokus pada kegiatan belajar, sehingga memiliki rata-rata prestasi akademik tertinggi (85-95).

Adapun visualisasi hasil K-Means Clustering pada Gambar di atas menunjukkan pengelompokan siswa ke dalam tiga cluster berbeda berdasarkan indikator penggunaan media sosial dan prestasi belajar. Setiap warna merepresentasikan kelompok siswa yang memiliki karakteristik serupa. Cluster_0 (oranye) menggambarkan kelompok dengan kecenderungan tertentu dalam penggunaan media sosial yang berbeda dari dua cluster lainnya. Cluster_1 (hijau) menunjukkan siswa dengan pola penggunaan media sosial yang sedang dan prestasi akademik menengah, sedangkan Cluster_2 (biru) menggambarkan kelompok siswa yang memiliki karakteristik berbeda dengan intensitas penggunaan tertentu serta rata-rata prestasi belajar yang juga bervariasi. Visualisasi ini memperjelas distribusi data serta pemisahan yang jelas antar cluster, sehingga memudahkan dalam memahami pola keterkaitan antara penggunaan media sosial dan prestasi belajar siswa.



Gambar 2. Visualisasi Hasil k-means cluster dalam penentuan prestasi belajar siswa

Gambar 2. Menampilkan hasil pengelompokan menggunakan algoritma K-Means Clustering dan divisualisasikan pada gambar di atas, terlihat bahwa data siswa terbagi secara jelas ke dalam tiga cluster yang berbeda. Hal ini menunjukkan bahwa terdapat perbedaan pola yang signifikan antara kelompok siswa dalam hal penggunaan media sosial serta hubungannya dengan prestasi belajar. Cluster yang terbentuk memberikan gambaran bahwa masing-masing kelompok memiliki kecenderungan perilaku dan capaian akademik yang khas. Dengan adanya hasil ini, peneliti dapat



lebih mudah mengidentifikasi kelompok siswa yang memerlukan perhatian khusus, baik dari segi bimbingan akademik maupun pengelolaan penggunaan media sosial, sehingga strategi peningkatan prestasi belajar dapat dilakukan secara lebih tepat sasaran.

3.3 Pengujian

Pengujian dilakukan untuk memastikan algoritma K-Means Clustering dapat berjalan sesuai tujuan penelitian dalam mengelompokkan siswa berdasarkan pola penggunaan media sosial dan prestasi belajar. Proses pengujian dilakukan menggunakan perangkat lunak RapidMiner, yang mempermudah dalam pengolahan data, pembentukan kluster, perhitungan nilai centroid, serta visualisasi hasil. Pada tahap ini ditentukan jumlah kluster (K) sebanyak tiga, yaitu kluster tinggi, sedang, dan rendah. Selanjutnya, sistem melakukan iterasi untuk menghitung jarak data terhadap centroid terdekat hingga terbentuk pengelompokan yang optimal. Berikut hasil pengujian hasil kinerja K-Means Cluster dalam penentuan prestasi siswa.

K-Means Clustering yang memetakan data ke dalam 3 cluster berdasarkan variable JHM (Jam Harian Media Sosial), berikut penjelasan gambar nya:

- Cluster 0, berisi siswa dengan JHM rendah–sedang (1–7 jam). Banyak siswa yang berada pada rentang 1–5 jam dengan warna merah–oranye (menandakan skor intensitas tinggi di variabel lain).
- Cluster 1, didominasi siswa dengan JHM menengah (4–6 jam), warnanya cenderung kuning (nilai moderat).
- Cluster 2, kelompok siswa dengan JHM tinggi (7–9 jam), ditandai warna biru (nilai lebih rendah pada indikator tertentu).

Berdasarkan hasil pengujian menggunakan algoritma K-Means Clustering dengan bantuan RapidMiner, diperoleh tiga kluster utama yang menggambarkan pola penggunaan media sosial siswa SMA Negeri 7 Sijunjung. Cluster 0 berisi kelompok siswa dengan intensitas penggunaan media sosial yang tinggi, yaitu rata-rata 7–9 jam per hari, yang cenderung berisiko mengalami penurunan fokus belajar. Cluster 1 merepresentasikan siswa dengan penggunaan media sosial pada tingkat sedang, yaitu 4–6 jam per hari, dengan pengaruh yang relatif moderat terhadap prestasi akademik. Sementara itu, Cluster 2 berisi siswa dengan penggunaan media sosial yang rendah hingga sedang, berkisar 1–5 jam per hari, yang umumnya lebih mampu menjaga konsentrasi dan menunjukkan kecenderungan prestasi belajar yang lebih baik. Grafik hasil klusterisasi menunjukkan distribusi siswa yang cukup jelas, sehingga memberikan gambaran bahwa semakin tinggi intensitas penggunaan media sosial, semakin besar kemungkinan berdampak pada penurunan capaian akademik.

4. KESIMPULAN

Penelitian ini membuktikan bahwa algoritma K-Means Clustering mampu mengelompokkan siswa SMA berdasarkan tingkat penggunaan media sosial dan prestasi belajar. Dari hasil pengujian menggunakan dataset siswa SMA Negeri 7 Sijunjung dengan 198 responden, diperoleh tiga kluster utama, yaitu C0 (Penggunaan Media Sosial Tinggi, memiliki prestasi rendah) yang beranggotakan 46 data sampel siswa, C1 (Penggunaan Media Sosial Sedang, Memiliki Prestasi Sedang) yang beranggotakan 80 sampel data siswa, dan C2 (Penggunaan Media Sosial Rendah, memiliki prestasi Tinggi) 72 sampel data. Hasil ini menunjukkan adanya perbedaan karakteristik pada masing-masing kluster, di mana intensitas penggunaan media sosial yang tinggi cenderung berkorelasi dengan penurunan konsentrasi dan prestasi belajar, sedangkan siswa dengan penggunaan media sosial yang rendah relatif memiliki capaian akademik yang lebih baik. Visualisasi kluster memperjelas distribusi data dan memudahkan dalam memahami pola keterkaitan antara penggunaan media sosial dan prestasi belajar. Dengan demikian, penelitian ini berkontribusi dalam memberikan pemahaman yang lebih mendalam mengenai dampak penggunaan media sosial terhadap capaian akademik siswa serta dapat menjadi dasar bagi sekolah, guru, maupun orang tua dalam menyusun strategi pembelajaran dan pengelolaan waktu yang lebih efektif. Adapun keterbatasan penelitian data hanya diperoleh dari satu sekolah (SMA Negeri 7 Sijunjung), Sehingga hasilnya belum tentu mewakili seluruh populasi siswa SMA, variabel yang digunakan masih terbatas pada intensitas penggunaan media sosial dan nilai akademik, belum mencakup factor lain seperti motivasi belajar, dukungan keluarga, atau kondisi psikologis siswa. Metode yang digunakan hanya K-Means Clustering, sehingga belum ada perbandingan dengan algoritma clustering lain yang mungkin memberikan hasil berbeda.

REFERENCES

- [1] D. Fitriah dan K. Khotimah, "Pengembangan Media Pembelajaran Kreatif Dan Interaktif Melalui Teknologi Informasi dan Komunikasi (TIK) Dalam Pendidikan," *IHSAN J. Pendidik. Islam*, vol. 3, hal. 1–8, 2025, doi: <https://doi.org/10.61104/ihsan.v3i1.433>.
- [2] E. S. Nomleni, P. A. Waluwandja, dan Y. Taebenu, "Dampak Penggunaan Media Sosial (Whatsapp) Terhadap Karakter Siswa Kelas VIII-I di SMP Negeri 1 Kota Kupang," vol. 8, no. 2, hal. 105–113, 2025, doi: <https://doi.org/10.70942/ciencias.v8i2.575>.
- [3] R. Anisa dan T. Nurhayati, "Dampak Penggunaan Media Sosial terhadap Perilaku Remaja," *J. Psikol. Pendidik. dan Konseling*, vol. 9, no. 1, hal. 25–34, 2023, doi: <https://doi.org/10.31004/jpion.v4i3.603>.
- [4] T. Lasmana dan A. Khoiriyati, "Analisis Dampak Pengguna Sosial Media Terhadap Produktivitas Akademik Mahasiswa," *J. Publ. Ilmu Manaj.*, vol. 4, no. 1, hal. 01–12, 2025, doi: <https://doi.org/10.55606/jupiman.v4i1.4828>.
- [5] L. A. A. D. A. ; S. Bahri, "Penggunaan media sosial sebagai," vol. 2, no. April, hal. 15–22, 2025.
- [6] D. Alsri Windra Doni, Aprizal Ponda, "Jurnal Pengabdian Masyarakat Cendikia Jenius," *J. Pengabd. Masy. Cendikia Jenius*, vol. 2, hal. 24–35, 2025, doi: <https://doi.org/10.70920/pengabmaskes.v2i2.89>.



- [7] S. Farid Raihan, "Analisis Pengaruh Media Sosial X Terhadap Sikap dan," *J. Indones. Manaj. Inform. dan Komun.*, vol. 5, no. 3, hal. 3138–3146, 2024, doi: <https://doi.org/10.35870/jimik.v5i3.991>.
- [8] M. Nur Afrilia, N. Rahaningsih, R. Danar Dana, dan N. Dienwati Nuris, "Optimasi Analisis Clustering Untuk Aktivitas Dan Respon Pengguna Media Sosial Dengan K-Means," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, hal. 148–155, 2024, doi: [10.36040/jati.v8i1.8334](https://doi.org/10.36040/jati.v8i1.8334).
- [9] S. Krisdianto Sitanggang, F. Rakhmat Umbara, dan H. Ashaury, "Klasifikasi Video Pada Media Sosial Youtube Dengan Menggunakan Metode K-Means Dan Support Vector Machine," *J. Locus Penelit. dan Pengabd.*, vol. 2, no. 10, hal. 1027–1032, 2023, doi: [10.58344/locus.v2i10.1732](https://doi.org/10.58344/locus.v2i10.1732).
- [10] N. Suarna, N. Rahaningsih, dan A. A. Suarna, "Optimalisasi Prestasi Akademik Siswa Melalui Pengelompokan Indeks Prestasi Dengan K-Means Clustering," *J. Kecerdasan Buatan dan Teknol. Inf.*, vol. 4, no. 2, hal. 198–207, 2025, doi: [10.69916/jkbti.v4i2.321](https://doi.org/10.69916/jkbti.v4i2.321).
- [11] D. Prasetyawan dan R. Gatra, "Analisis Cluster untuk Pengelompokan Kemampuan Penguasaan ICT Menggunakan K-Means dan Autoencoder," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 10, no. 2, hal. 145–157, 2025, doi: [10.14421/jiska.2025.10.2.145-157](https://doi.org/10.14421/jiska.2025.10.2.145-157).
- [12] D. Helmi Saputra, Rivaldo Aldino Uneputty, "Analisis Pola Kegiatan Belajar Mahasiswa Terhadap Keberhasilan Akademik Menggunakan Algoritma Apriori," *J. sains Inform. Terap. (JSIT)*, vol. 2, no. 1, hal. 16–20, 2023, doi: <https://doi.org/10.62357/jsit.v4i2.567>.
- [13] W. I. Permatasari dan V. Tundjungsari, "Implementasi Algoritma K-Means untuk Mengetahui Minat Siswa SMA Terhadap Mata Pelajaran Teknologi Informasi dan Komunikasi," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 1, hal. 168–174, 2024, doi: [10.57152/malcom.v5i1.1639](https://doi.org/10.57152/malcom.v5i1.1639).
- [14] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *J. Ilm. Inform. dan Ilmu Komput.*, vol. 3, no. 1, hal. 46–56, 2024, doi: [10.58602/jima-ilkom.v3i1.26](https://doi.org/10.58602/jima-ilkom.v3i1.26).
- [15] N. A. Y. -, L. E. B. -, G. C. H. R. -, M. F. Z. -, A. -, dan F. R. -, "Analisis Perbandingan K-Means Dan Dbscan Dalam Pengelompokan Data Travel Review Ratings Menggunakan Evaluasi Silhouette Index Dan Davies-Bouldin Index," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 3, 2025, doi: [10.23960/jitet.v13i3.6884](https://doi.org/10.23960/jitet.v13i3.6884).
- [16] S. S. Putro, M. Syarief, dan E. M. S. Rochman, "Penerapan Algoritma K-Means, Dbscan, Dan Ahc Untuk Clustering Kualitas Garam Pada Pt. Garam (Persero)," *Multitek Indones.*, vol. 18, no. 2, hal. 114–125, 2025, doi: [10.24269/mtkind.v18i2.8501](https://doi.org/10.24269/mtkind.v18i2.8501).
- [17] S. F. Harianti, R. A. Dewanto, R. Rahmiati, dan A. Suharso, "PENGARUH PENGGUNAAN MEDIA SOSIAL TERHADAP PRESTASI BELAJAR SISWA SMA N 3 DEPOK," *J. Elem. Sch.*, vol. 6, no. 2, hal. 166–173, 2023, doi: [10.31004/jrpp.v6i1.16072](https://doi.org/10.31004/jrpp.v6i1.16072).
- [18] A. B. Fatika dan S. Suyato, "Pengaruh Media Sosial Terhadap Pencapaian Prestasi Belajar Pendidikan Pancasila di SMA Muhammadiyah 1 Klaten," *Agora*, vol. 13, no. 5, hal. 512–521, 2025, doi: [10.21831/agora.v13i5.23116](https://doi.org/10.21831/agora.v13i5.23116).
- [19] I. Q. Azizah dan I. Anshori, "Pengaruh Penggunaan Tiktok Sebagai Media Sosial Terhadap Fokus Belajar Dan Produktivitas Mahasiswa," *JASIMA J. Komun. Korporasi Dan Media*, vol. 6, no. 1, hal. 88–100, 2025, doi: <https://doi.org/10.30872/jasima.v6i1.2425>.
- [20] M. F. Akmal, M. F. N. G. Ratumbusang, M. Hasanah, dan B. Nor, "Pengaruh Media Sosial Terhadap Motivasi Belajar Mahasiswa Pendidikan Ekonomi FKIP ULM," *J. Pendidik. Ekon.*, vol. 12, no. 2, hal. 168–175, 2024, doi: [10.26740/jupe.v12n2.p168-175](https://doi.org/10.26740/jupe.v12n2.p168-175).
- [21] N. Musdira, "Pengaruh Penggunaan Media Sosial Terhadap Prestasi Akademik Peserta Didik Di SMK Negeri 2 Gowa," vol. 03, hal. 66–74, 2025, [Daring]. Tersedia pada: <https://journal.diginus.id/index.php/JUPITER/index>
- [22] C. Paramita, F. A. Rafrastara, dan C. Supriyanto, "Pemanfaatan Algoritma K-Means untuk Membuktikan Implementasi Undang-Undang Pelanggaran Hukum Korupsi di Pengadilan Negeri Banjarmasin," *J. Inform. J. Pengemb. IT*, vol. 8, no. 2, hal. 149–154, 2023, doi: [10.30591/jpit.v8i2.5216](https://doi.org/10.30591/jpit.v8i2.5216).
- [23] R. M. Koretsky, "Python3," *Raspberry Pi OS Syst. Adm. with Syst. Python*, vol. 8, no. 1, hal. 175–305, 2023, doi: [10.1201/b23421-3](https://doi.org/10.1201/b23421-3).