

# Implementasi Data Mining Dalam Pengelompokan Perilaku Penggunaan Media Sosial Siswa SMA Menerapkan Algoritma K-Means

Utari\*, Muhammad Iqbal

Fakultas Pasca Sarjana, Program Studi Magister Teknologi Informasi, Universitas Pembangunan Panca Budi, Medan, Indonesia

Email: <sup>1,\*</sup>utarijaya1999@gmail.com, <sup>2</sup>muhammadiqbal@gmail.com

Email Correspondencial: utarijaya1999@gmail.com

**Abstrak**—Perkembangan teknologi informasi dan komunikasi telah mendorong meningkatnya penggunaan media sosial di kalangan remaja, khususnya siswa Sekolah Menengah Atas (SMA). Media sosial tidak hanya dimanfaatkan sebagai sarana komunikasi dan hiburan, tetapi juga berpotensi memengaruhi perilaku belajar siswa. Penggunaan yang berlebihan dapat menurunkan konsentrasi dan efektivitas belajar, sehingga diperlukan analisis untuk mengidentifikasi pola perilaku penggunaan media sosial secara sistematis. Penelitian ini bertujuan untuk mengelompokkan perilaku penggunaan media sosial siswa SMA menggunakan algoritma K-Means Clustering. Data diperoleh melalui kuesioner tertutup dengan skala numerik yang mencakup tiga variabel, yaitu frekuensi penggunaan media sosial, durasi penggunaan, dan jumlah platform media sosial yang digunakan. Sebanyak 20 data sampel digunakan sebagai ilustrasi proses perhitungan. Tahapan analisis meliputi normalisasi data menggunakan metode Min–Max, penentuan jumlah kluster sebanyak tiga kluster ( $K = 3$ ), perhitungan jarak menggunakan Euclidean Distance, serta pembaruan centroid secara iteratif. Berdasarkan hasil perhitungan algoritma K-Means hingga mencapai kondisi konvergen, perilaku penggunaan media sosial siswa SMA berhasil dikelompokkan ke dalam tiga cluster, yaitu penggunaan rendah dengan centroid (3,50; 1,25; 2,75), penggunaan sedang dengan centroid (7,25; 3,25; 5,88), dan penggunaan tinggi dengan centroid (12,25; 5,75; 8,88). Hasil ini menunjukkan bahwa algoritma K-Means efektif dalam mengelompokkan perilaku penggunaan media sosial siswa berdasarkan perbedaan frekuensi akses, durasi penggunaan, dan tingkat interaksi media sosial.

**Kata kunci:** Media Sosial; K-Means Clustering; Data Mining; Siswa SMA

**Abstract**—The development of information and communication technology has encouraged the increasing use of social media among teenagers, especially high school students (SMA). Social media is not only used as a means of communication and entertainment, but also has the potential to influence students' learning behavior. Excessive use can reduce concentration and learning effectiveness, so analysis is needed to systematically identify patterns of social media usage behavior. This study aims to group the social media usage behavior of high school students using the K-Means Clustering algorithm. Data were obtained through a closed questionnaire with a numerical scale that includes three variables, namely the frequency of social media use, duration of use, and the number of social media platforms used. A total of 20 sample data were used to illustrate the calculation process. The analysis stages include data normalization using the Min-Max method, determining the number of clusters as many as three clusters ( $K = 3$ ), calculating distances using Euclidean Distance, and iteratively updating the centroids. Based on the results of the K-Means algorithm calculations until reaching a convergent condition, the social media usage behavior of high school students was successfully grouped into three clusters, namely low usage with centroids (3.50; 1.25; 2.75), moderate usage with centroids (7.25; 3.25; 5.88), and high usage with centroids (12.25; 5.75; 8.88). These results indicate that the K-Means algorithm is effective in grouping students' social media usage behavior based on differences in access frequency, duration of use, and level of social media interaction.

**Keywords:** Social Media; K-Means Clustering; Data Mining; Siswa SMA

## 1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah membawa perubahan signifikan dalam kehidupan manusia, termasuk dalam dunia Pendidikan. 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. Namun, penggunaan media sosial yang tidak terkontrol berpotensi mempengaruhi prestasi belajar siswa[1].

Perilaku penggunaan media sosial siswa SMA menunjukkan karakteristik yang beragam, baik dari segi frekuensi, durasi, maupun tingkat interaksi. Perbedaan tersebut menyebabkan kesulitan dalam mengidentifikasi kelompok siswa dengan tingkat penggunaan media sosial tertentu jika hanya mengandalkan pengamatan subjektif. Selama ini, analisis perilaku penggunaan media sosial siswa umumnya dilakukan secara deskriptif tanpa pengelompokan yang jelas, sehingga hasil yang diperoleh kurang optimal untuk dijadikan dasar dalam pengambilan keputusan. Kondisi ini menunjukkan perlunya pendekatan berbasis data yang mampu mengelompokkan siswa secara objektif berdasarkan kesamaan karakteristik perilaku penggunaan media sosial[2][3]. Objek penelitian dalam penelitian ini adalah siswa Sekolah Menengah Atas (SMA) yang aktif menggunakan media sosial. Data penelitian berupa perilaku penggunaan media sosial siswa, yang diperoleh melalui kuesioner tertutup dengan skala numerik, meliputi indikator frekuensi penggunaan, durasi penggunaan, serta tingkat interaksi di media sosial[4].

Namun, beberapa permasalahan muncul dari besarnya volume data (Big Data) yang terus dihasilkan setiap detiknya diseluruh platform. Permasalahan tersebut diantaranya belum adanya pengelompokan perilaku penggunaan media sosial siswa SMA secara sistematis dan berbasis data[5][6]. Kesulitan dalam mengidentifikasi siswa dengan tingkat penggunaan media sosial rendah, sedang, dan tinggi. Minimnya pemanfaatan teknik data mining, khususnya metode clustering, dalam menganalisis perilaku penggunaan media sosial siswa di tingkat SMA, maka solusi atas permasalahan tersebut, penelitian ini menerapkan algoritma K-Means Clustering untuk mengelompokkan perilaku penggunaan media

sosial siswa SMA. Algoritma K-Means dipilih karena kemampuannya dalam mengelompokkan data ke dalam sejumlah cluster berdasarkan tingkat kemiripan karakteristik[7][8][9]. Dalam penelitian ini, data perilaku siswa dikelompokkan ke dalam tiga cluster, yaitu penggunaan rendah, penggunaan sedang, dan penggunaan tinggi. Hasil pengelompokan ini diharapkan dapat memberikan gambaran yang lebih jelas mengenai pola penggunaan media sosial siswa serta menjadi dasar bagi pihak sekolah dalam merancang strategi pembinaan, pengawasan, dan kebijakan penggunaan media sosial yang lebih efektif.

Data mining dapat dipahami sebagai suatu proses penemuan pengetahuan (*knowledge discovery*) atau informasi bernilai dari kumpulan data berukuran besar melalui penerapan berbagai algoritma yang telah teruji secara ilmiah. Dalam penelitian ini, algoritma yang digunakan adalah K-Means Clustering. K-Means merupakan metode klusterisasi non-hierarkis yang berfungsi mengelompokkan data berdasarkan tingkat kesamaan karakteristik intrinsik antar *data point*[10]. Konsep dasar algoritma ini adalah membagi sejumlah objek data ke dalam kluster, di mana setiap objek akan ditempatkan pada kluster yang memiliki nilai rata-rata atau pusat kluster (*centroid*) terdekat. Pemilihan algoritma K-Means dalam penelitian ini didasarkan pada beberapa pertimbangan. Pertama, K-Means memiliki tingkat efisiensi komputasi yang tinggi sehingga mampu menangani data interaksi media sosial dalam jumlah besar secara efektif. Kedua, algoritma ini mampu menyederhanakan pola perilaku pengguna yang kompleks ke dalam sejumlah kluster yang bersifat diskrit dan mudah diinterpretasikan[11][12]. Dengan karakteristik tersebut, K-Means dinilai sangat sesuai untuk mencapai tujuan utama penelitian, yaitu mengelompokkan perilaku penggunaan media Sosial guna memperoleh segmentasi perilaku penggunaan media sosial secara lebih terstruktur[13][14].

Penerapan teknik klusterisasi dalam kajian perilaku pengguna media sosial telah banyak dilakukan sebelumnya. Berbagai penelitian dalam beberapa tahun terakhir menunjukkan adanya peningkatan pemanfaatan metode data mining dalam analisis fenomena sosial. Sebagai contoh, penelitian oleh Muhamad Ilhan Mansiz & Zaehol Fatah (2024) menggunakan K-Means Clustering untuk mengelompokkan pengguna media sosial berdasarkan pola interaksi mereka, seperti frekuensi kunjungan, durasi interaksi, serta perilaku kunjungan berdasarkan waktu dan lokasi. Variabel-variabel ini diolah melalui normalisasi dan pengkodean sehingga dapat dianalisis dengan algoritma K-Means. Hasil clustering mengidentifikasi dua kelompok utama pengguna dengan karakteristik keterlibatan yang berbeda. Penelitian ini memberikan gambaran aplikatif tentang bagaimana K-Means dapat digunakan untuk memahami perilaku pengguna media sosial lebih mendalam.

Selanjutnya penelitian oleh Rahmah Dila et al (2025) mempelajari hubungan antara penggunaan media sosial dan prestasi belajar siswa SMA. Penelitian ini menargetkan data yang mencakup durasi penggunaan media sosial, waktu aktif, tingkat kecanduan, serta rata-rata nilai akademik. K-Means Clustering digunakan untuk mengelompokkan siswa ke dalam cluster dengan karakteristik penggunaan media sosial tinggi, sedang, dan rendah, serta melihat keterkaitannya dengan prestasi akademik. Hasil penelitian menunjukkan bahwa algoritma K-Means efektif mengelompokkan siswa berdasarkan pola penggunaan serta memberikan informasi yang bermanfaat tentang pengaruh media sosial terhadap prestasi belajar secara empiris. Penelitian oleh Farid Raihan & Sutisna (2024) yang menganalisis pola penggunaan media sosial *Platform X* dan hubungannya dengan sikap serta perilaku remaja. Data dikumpulkan melalui survei yang mencakup frekuensi penggunaan, jenis konten yang dikonsumsi, serta interaksi sosial di platform tersebut. Algoritma K-Means Clustering digunakan untuk mengelompokkan remaja berdasarkan pola penggunaan dan karakteristik perilaku mereka. Hasilnya menunjukkan adanya beberapa cluster yang menggambarkan perbedaan intensitas penggunaan media sosial serta pengaruhnya terhadap perilaku remaja, memberikan wawasan empiris tentang dampak media sosial pada perkembangan sikap anak muda.

Meskipun penelitian-penelitian sebelumnya telah meletakkan dasar yang kuat dalam klusterisasi aktivitas, Kontribusi fundamental penelitian ini terletak pada upaya untuk menghasilkan pemetaan aktivitas online yang lebih komprehensif dan multidimensi. Alih-alih hanya berfokus pada satu dimensi, penelitian ini berupaya mengintegrasikan berbagai metrik interaksi untuk membangun profil interaksi yang lebih lengkap. Melalui penerapan algoritma K-Means yang terstruktur pada fitur interaksi yang kaya ini, penelitian ini bertujuan memberikan gambaran segmentasi yang tajam tentang tipe pengguna berdasarkan tingkat dan jenis aktivitas mereka.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan penelitian

Penelitian ini dengan sistematis melalui beberapa tahapan yang banyak agar memastikan tercapainya tujuan utama, yaitu Perilaku penggunaan media sosial. Ada pun kerangka kerja penelitian ini di mulain dari identifikasi permasalahan yang jelas sehingga kesimpulan dari hasil karakteristik data yang diperoleh. Berikut tahapan dari penelitian yang dilakukan penulis:

#### a. Identifikasi Masalah

Tahap awal penelitian ini bertujuan untuk menetapkan fokus serta batasan kajian yang akan dilakukan. Pada tahap ini dilakukan identifikasi terhadap kesenjangan pengetahuan maupun permasalahan yang perlu diselesaikan. Dalam penelitian ini, permasalahan utama yang dihadapi adalah besarnya jumlah dan kompleksitas data perilaku penggunaan media sosial, yang menyebabkan sulitnya mengenali pola perilaku secara manual. Oleh karena itu, diperlukan suatu pendekatan komputasional yang mampu mengolah dan menganalisis data dalam jumlah besar secara efektif dan efisien. Pendekatan *data mining*, khususnya teknik *clustering*, dipandang sebagai solusi yang tepat untuk

mengelompokkan perilaku pengguna media sosial berdasarkan tingkat kemiripan karakteristik data. Dengan adanya proses pengelompokan ini, diharapkan pola perilaku penggunaan media sosial dapat diidentifikasi secara lebih sistematis dan objektif, sehingga dapat memberikan informasi yang bernilai sebagai dasar analisis dan pengambilan keputusan dalam konteks pendidikan.

b. Pengumpulan Data

Data penelitian ini diperoleh melalui penyebaran kuesioner tertutup yang dirancang menggunakan skala numerik. Kuesioner tertutup dipilih karena mampu menghasilkan data kuantitatif yang sistematis, terukur, dan seragam antarresponden, sehingga sesuai untuk dianalisis menggunakan metode data mining, khususnya algoritma K-Means Clustering. Data yang diperoleh dari hasil pengisian kuesioner selanjutnya dikompilasi dan diolah dalam bentuk data numerik, sehingga dapat dianalisis secara komputasional. Karakteristik data tersebut dinilai sesuai untuk diterapkan metode *data mining*, khususnya algoritma K-Means Clustering, karena algoritma ini bekerja secara efektif pada data kuantitatif dan bertujuan untuk mengelompokkan objek penelitian berdasarkan tingkat kemiripan nilai atribut.

c. Studi Literatur

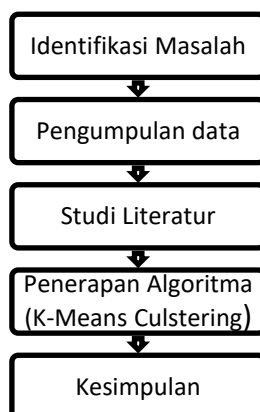
Tahap studi literatur dilakukan untuk membangun landasan teori yang kuat sebagai dasar penelitian. Peneliti mempelajari berbagai referensi yang relevan, seperti konsep perilaku penggunaan media sosial, *data mining*, teknik *clustering*, serta algoritma K-Means. Studi literatur ini bertujuan untuk memahami metode yang digunakan, menemukan celah penelitian sebelumnya, serta memastikan bahwa penelitian yang dilakukan memiliki kontribusi ilmiah yang jelas dan berbeda. Melalui studi literatur, peneliti memperoleh pemahaman yang lebih mendalam mengenai karakteristik data, tahapan penerapan algoritma K-Means, serta kelebihan dan keterbatasan metode tersebut dalam mengelompokkan data berdasarkan tingkat kemiripan. Selain itu, kajian terhadap penelitian-penelitian terdahulu dilakukan untuk mengidentifikasi pendekatan yang telah digunakan, hasil yang telah dicapai, serta permasalahan atau keterbatasan yang masih belum terpecahkan.

d. Penerapan Algoritma (K-Means Clustering)

Tahapan utama dalam penelitian ini adalah pelaksanaan metode analisis yang telah ditetapkan. Setelah data melalui proses persiapan dan siap untuk dianalisis, algoritma K-Means digunakan sebagai teknik pengelompokan. Tahap ini meliputi penentuan jumlah kluster (K) yang paling sesuai, penetapan pusat kluster awal (centroid), serta proses iteratif untuk mengelompokkan pengguna berdasarkan tingkat kesamaan pola interaksi yang dimiliki. Output dari proses ini berupa terbentuknya beberapa kelompok (kluster) pengguna yang mencerminkan variasi karakteristik aktivitas daring yang berbeda.

e. Kesimpulan

Tahap akhir penelitian ini berfokus pada penafsiran serta perumusan hasil yang diperoleh dari proses klusterisasi. Kesimpulan disusun untuk menjawab tujuan penelitian yang telah ditetapkan sejak awal, sekaligus merangkum temuan utama penelitian, termasuk karakteristik masing-masing kluster pengguna yang berhasil diidentifikasi. Selain itu, pada tahap ini juga disajikan rekomendasi atau implikasi praktis yang dapat dimanfaatkan sebagai bahan pertimbangan dalam pengembangan platform maupun dalam kajian lanjutan terkait analisis perilaku digital di masa yang akan datang. Hasil penelitian menunjukkan bahwa algoritma K-Means mampu diterapkan secara efektif dalam mengelompokkan data perilaku penggunaan media sosial siswa ke dalam tiga kluster utama, yaitu kluster dengan intensitas penggunaan rendah, sedang, dan tinggi. Maka dari itu alur tahapan penelitian disajikan pada Gambar 1 sebagai berikut:



Gambar 1. Tahapan penelitian

## 2.2 Algoritma K-Means

Algoritma K-Means merupakan salah satu teknik clustering dalam bidang data mining yang berfungsi untuk mengelompokkan sekumpulan data ke dalam sejumlah kluster berdasarkan tingkat kesamaan karakteristik antar data. Mekanisme kerja algoritma ini dilakukan secara iteratif dengan menentukan pusat kluster awal (centroid), kemudian setiap data dialokasikan ke kluster yang memiliki jarak terdekat terhadap centroid tersebut. Setelah proses pengelompokan dilakukan, posisi centroid diperbarui berdasarkan nilai rata-rata dari seluruh data yang berada dalam masing-masing

klaster[15]. Proses ini terus diulang hingga posisi centroid mencapai kondisi konvergen atau tidak mengalami perubahan yang signifikan. Algoritma K-Means dinilai efektif dalam mengidentifikasi pola-pola tersembunyi pada data berskala besar serta membantu mengorganisasi objek atau perilaku ke dalam kelompok yang lebih terstruktur. Oleh karena itu, algoritma ini banyak diterapkan dalam berbagai bidang, seperti analisis perilaku pengguna, segmentasi pelanggan, kajian pola interaksi, hingga pengelompokan teks pada platform media sosial[16]. Algoritma K-Means menggunakan perhitungan jarak Euclidean untuk menentukan kedekatan data dengan pusat cluster, yang dirumuskan sebagai berikut:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan variabel pada rumus tersebut menjelaskan bahwa  $d(x,y)$  merupakan jarak antara suatu data dengan pusat cluster, di mana  $x_i$  adalah nilai atribut ke- $i$  dari data yang dihitung, sedangkan  $y_i$  adalah nilai atribut ke- $i$  dari centroid, dengan  $n$  menyatakan jumlah atribut atau variabel yang digunakan dalam perhitungan,  $\epsilon$  menunjukkan operasi penjumlahan dari atribut ke-1 hingga ke- $n$ , dan simbol  $\sqrt{\quad}$  menyatakan akar kuadrat dari hasil perhitungan tersebut.

### 2.3 Data Mining

Data mining merupakan suatu proses eksplorasi dan analisis terhadap data dalam jumlah besar yang bertujuan untuk mengidentifikasi pola, keterkaitan, kecenderungan, maupun pengetahuan baru yang sebelumnya belum terungkap secara langsung. Proses ini dilakukan dengan memanfaatkan berbagai pendekatan, seperti metode statistik, kecerdasan buatan, dan pembelajaran mesin, guna mengekstraksi informasi yang bernilai sebagai dasar dalam pengambilan keputusan. Melalui data mining, pengolahan data dapat dilakukan secara otomatis dan terstruktur, sehingga memungkinkan ditemukannya pola-pola tersembunyi yang sulit diidentifikasi menggunakan teknik analisis konvensional[17]. Dalam konteks penelitian, data mining berperan penting dalam memahami karakteristik data serta dalam melakukan prediksi atau pengelompokan objek berdasarkan kesamaan atribut tertentu. Data mining tidak hanya berfokus pada pengolahan data, tetapi juga mencakup tahapan pemilihan data, pembersihan data (*data cleaning*), transformasi data, penerapan algoritma analisis, serta evaluasi hasil yang diperoleh. Melalui tahapan tersebut, data mining mampu mengidentifikasi pola keteraturan, tren, anomali, maupun hubungan antar variabel yang sebelumnya tidak diketahui secara eksplisit. Oleh karena itu, data mining sering dipandang sebagai bagian integral dari proses *Knowledge Discovery in Databases* (KDD), yaitu proses penemuan pengetahuan baru dari basis data[18].

### 2.3 Media Sosial

Media sosial merupakan platform berbasis teknologi digital yang memungkinkan pengguna untuk membuat, berbagi, serta berinteraksi melalui berbagai bentuk konten, seperti teks, gambar, audio, dan video, dalam jaringan daring. Dalam konteks pendidikan, khususnya pada siswa Sekolah Menengah Atas (SMA), media sosial tidak hanya dimanfaatkan sebagai sarana komunikasi dan hiburan, tetapi juga berpotensi memengaruhi proses belajar, interaksi sosial, serta pembentukan perilaku dan sikap. Intensitas dan pola penggunaan media sosial yang beragam dapat menimbulkan dampak positif maupun negatif, sehingga diperlukan kajian ilmiah untuk menganalisis dan mengelompokkan perilaku penggunaan media sosial secara sistematis[19]. Media sosial memfasilitasi komunikasi dua arah secara cepat dan luas, sehingga pengguna dapat berpartisipasi aktif dalam pertukaran informasi, pembentukan komunitas, serta pengembangan jejaring sosial tanpa dibatasi oleh ruang dan waktu. Karakteristik utama media sosial terletak pada sifatnya yang interaktif, partisipatif, dan kolaboratif, di mana pengguna tidak hanya berperan sebagai penerima informasi, tetapi juga sebagai produsen konten (*user-generated content*). Melalui mekanisme jejaring sosial digital, media sosial memungkinkan terbentuknya hubungan sosial, komunitas virtual, serta pola komunikasi dua arah yang bersifat dinamis dan berkelanjutan. Selain itu, media sosial juga mendukung penyebaran informasi secara cepat dan luas, sehingga memiliki pengaruh signifikan terhadap pola komunikasi, perilaku, dan interaksi sosial penggunanya[20].

## 3. HASIL DAN PEMBAHASAN

### 3.1 Data Sampel

Data sampel yang digunakan dalam penelitian ini merupakan data perilaku penggunaan media sosial siswa Sekolah Menengah Atas (SMA) yang diperoleh melalui hasil pengisian kuesioner tertutup dengan menggunakan skala numerik. Data tersebut mencerminkan karakteristik aktivitas siswa dalam menggunakan media sosial, yang meliputi beberapa indikator perilaku, seperti frekuensi akses, durasi penggunaan, serta tingkat interaksi dalam berbagai platform media sosial. Penggunaan skala numerik bertujuan untuk mengubah data kualitatif menjadi bentuk kuantitatif sehingga dapat diolah dan dianalisis secara komputasional. Berikut pada Tabel 1 menunjukkan variabel yg digunakan dalam penelitian ini.

Tabel 1. Variabel

Siswa	X1	X2	X3
S1	3	1	2
S2	4	1	3
S3	2	1	2

Siswa	X1	X2	X3
S4	5	2	4
S5	6	2	5
S6	7	3	6
S7	8	3	7
S8	9	4	7
S9	10	4	8
S10	11	5	8
S11	12	5	9
S12	13	6	9
S13	14	6	10
S14	15	7	10
S15	5	2	4
S16	6	3	5
S17	8	4	6
S18	9	5	7
S19	11	6	8
S20	12	7	9

### 3.2 Penerapan Algoritma K-Means

Ada pun tahapan algoritma K-Means ialah :

- Menentukan Jumlah Cluster (k)  
k = 3, dengan asumsi:  
Cluster 1 → Penggunaan rendah  
Cluster 2 → Penggunaan sedang  
Cluster 3 → Penggunaan tinggi
- Menentukan Titik Pusat Awal (Centroid Awal)  
Dipilih secara representatif dari data:  
C1 (Rendah) = S3 → (2, 1, 2)  
C2 (Sedang) = S7 → (8, 3, 7)  
C3 (Tinggi) = S14 → (15, 7, 10)
- Menghitung Jarak ke Centroid  
Rumus Jarak Euclidean

$$D = \sqrt{(X1 - C1)^2 + (X2 - C2)^2 + (X3 - C3)^2}$$

Perhitungan

S1 (3,1,2)

$$\text{Ke C1} = \sqrt{(3-2)^2 + (1-1)^2 + (2-2)^2} = \sqrt{1} = 1,00$$

$$\text{Ke C2} = \sqrt{(3-8)^2 + (1-3)^2 + (2-7)^2} = \sqrt{25 + 4 + 25} = \sqrt{25} = 7,35$$

$$\text{Ke C3} = \sqrt{(3-15)^2 + (1-7)^2 + (2-10)^2} = \sqrt{144 + 36 + 64} = \sqrt{244} = 15,62$$

Maka didapatkan hasil Perhitungan jarak dan cluster dapat dilihat pada tabel 2.

**Tabel 2.** Jarak dan Cluster

Siswa	d(C1)	d(C2)	d(C3)	Cluster
S1	1,00	7,35	15,62	C1
S2	2,24	6,48	14,87	C1
S3	0,00	7,07	15,26	C1
S4	3,74	4,58	12,53	C1
S5	5,10	3,32	10,77	C2
S6	6,48	1,41	9,11	C2
S7	7,07	0,00	7,62	C2
S8	8,60	1,41	6,16	C2
S9	9,90	2,45	5,10	C3
S14	15,26	7,62	0,00	C3

- Pengelompokan Data Berdasarkan Jarak Terdekat ke Centroid  
Berdasarkan hasil perhitungan jarak Euclidean pada iterasi pertama, setiap data dikelompokkan ke dalam cluster dengan jarak terpendek (minimum) terhadap centroid awal. Berikut perhitungan dari hasil pengelompokan iterasi Pertama :  
1. Cluster 1 (Penggunaan Rendah – C1)  
2. Cluster 2 (Penggunaan Sedang – C2)

## 3. Cluster 3 (Penggunaan Tinggi – C3)

## e. Memperbarui Nilai Centroid Baru

Centroid baru diperoleh dari rata-rata setiap variabel (X1, X2, X3) pada masing-masing cluster.berikut Perhitungan Centroid Baru Cluster 1 :

$$X1 = \frac{3+4+2+5}{4} = \frac{14}{4} = 3,50$$

$$X2 = \frac{1+1+1+4}{4} = \frac{4}{4} = 1,25$$

$$X3 = \frac{2+3+2+4}{4} = \frac{11}{4} = 2,75$$

Maka hasil dari centroid C1 adalah : (3,50 , 1,25 , 2,75)

Perhitungan Centroid Baru Cluster 2

$$X1 = \frac{6+7+8+9+5+6+8=9}{8} = \frac{58}{8} = 7,25$$

$$X2 = \frac{2+3+3+4+2+3+4+5}{8} = \frac{26}{8} = 3,25$$

$$X3 = \frac{5+6+7+7+4+5+6+7}{8} = \frac{47}{8} = 5,88$$

Maka hasil dari centroid C1 adalah : (7,25 , 3,25 , 5,88)

Perhitungan Centroid Baru Cluster 3

$$X1 = \frac{10+11+12+13+14+15+11+12}{8} = \frac{98}{8} = 12,25$$

$$X2 = \frac{4+5+6+6+7+6+7}{8} = \frac{46}{8} = 5,75$$

$$X3 = \frac{8+8+9+9+10+10+8+9}{8} = \frac{71}{8} = 8,88$$

Maka hasil dari centroid C1 adalah : (12,25 , 5,75 , 8,88)

Berikut pada tabel 3 merupakan hasil dari centeroid baru seperti tabel dibawa:

**Tabel 3.** Centroid Baru

Cluster	Centroid Awal	Centroid Baru
C1	(2,1,2)	(3,50 , 1,25 , 2,75)
C2	(8,3,7)	(7,25 , 3,25 , 5,88)
C3	(15,7,10)	(12,25 , 5,75 , 8,88)

## f. Pengulangan Perhitungan Jarak (Iterasi Berikutnya)

Setelah diperoleh centroid baru pada iterasi pertama, selanjutnya dilakukan pengulangan langkah ke-3, yaitu menghitung kembali jarak setiap data terhadap centroid terbaru. Tujuan dari langkah ini adalah untuk mengetahui apakah terjadi perpindahan keanggotaan cluster pada setiap data.

Centroid baru yang digunakan pada iterasi kedua adalah sebagai berikut:

$$C1 = (3,50 ; 1,25 ; 2,75)$$

$$C2 = (7,25 ; 3,25 ; 5,88)$$

$$C3 = (12,25 ; 5,75 ; 8,88)$$

Perhitungan jarak tetap menggunakan rumus Euclidean:

$$D = \sqrt{(X1 - C1)^2 + (X2 - C2)^2 + (X3 - C3)^2}$$

Setiap data dihitung jaraknya ke ketiga centroid baru tersebut, kemudian dikelompokkan ke dalam cluster dengan jarak terpendek.

Perhitungan Iterasi Kedua

Data S1 (3, 1, 2)

$$Ke C1 = \sqrt{(3 - 3,50)^2 + (1 - 1,25)^2 + (2 - 2,75)^2} = \sqrt{0,25 + 0,06 + 0,56} = \sqrt{0,87} = 0,93$$

$$Ke C2 = \sqrt{(3 - 7,25)^2 + (1 - 3,25)^2 + (2 - 5,88)^2} = \sqrt{18,06 + 5,06 + 15,05} = \sqrt{38,17} = 6,18$$

$$Ke C3 = \sqrt{(3 - 12,25)^2 + (1 - 5,75)^2 + (2 - 8,88)^2} = \sqrt{88,56 + 22,56 + 47,33} = \sqrt{158,45} = 12,47$$

S1 tetap berada pada Cluster 1

Data S5 (6, 2, 5)

$$Ke C1 = \sqrt{(6 - 3,50)^2 + (2 - 1,25)^2 + (5 - 2,75)^2} = \sqrt{6,25 + 0,56 + 5,06} = \sqrt{11,87} = 3,45$$

$$Ke C2 = \sqrt{(6 - 7,25)^2 + (2 - 3,25)^2 + (5 - 5,88)^2} = \sqrt{1,56 + 1,56 + 0,77} = \sqrt{3,89} = 1,97$$

$$Ke C3 = \sqrt{(6 - 12,25)^2 + (2 - 5,75)^2 + (5 - 8,00)^2} = \sqrt{39,06 + 14,06 + 15,05} = \sqrt{68,17} = 8,26$$

S5 tetap berada pada Cluster 2

Data S10 (11, 5, 8)

$$Ke C1' = 10,23$$

$$Ke C2' = 4,43$$

$$Ke C3' = 1,98$$

S10 tetap berada pada Cluster 3.

Berdasarkan perhitungan iterasi kedua, terlihat bahwa proses pengelompokan data telah menunjukkan kestabilan karena tidak terjadi perpindahan keanggotaan cluster pada data yang diuji. Data S1 tetap memiliki jarak terpendek terhadap centroid C1, data S5 paling dekat dengan centroid C2, dan data S10 tetap berada pada cluster C3, yang menunjukkan bahwa centroid baru yang terbentuk pada iterasi pertama sudah merepresentasikan karakteristik masing-masing kelompok data dengan baik. Kondisi ini mengindikasikan bahwa proses iterasi mulai menuju konvergensi, di mana perubahan centroid tidak lagi menyebabkan perubahan signifikan pada hasil pengelompokan, sehingga cluster yang terbentuk semakin optimal dan konsisten dalam merepresentasikan pola data. Tabel 4 merupakan hasil iterasi kedua yang diperoleh:

Tabel 4. Hasil Iterasi Kedua

Cluster	Anggota
C1	S1, S2, S3, S4
C2	S5, S6, S7, S8, S15, S16, S17, S18
C3	S9, S10, S11, S12, S13, S14, S19, S20

Hasil perhitungan jarak pada iterasi kedua menunjukkan bahwa keanggotaan setiap cluster tidak mengalami perubahan. Dengan demikian Proses pengelompokan menggunakan algoritma K-Means telah mencapai kondisi konvergen, Iterasi dihentikan pada iterasi kedua, Hasil cluster yang diperoleh dinyatakan stabil dan final.

#### 4. KESIMPULAN

Berdasarkan hasil perhitungan algoritma K-Means yang dilakukan hingga mencapai kondisi konvergen, diperoleh tiga cluster perilaku penggunaan media sosial siswa SMA, yaitu cluster penggunaan rendah, sedang, dan tinggi. Cluster penggunaan rendah memiliki centroid akhir sebesar (3,50; 1,25; 2,75) yang menunjukkan frekuensi akses, durasi penggunaan, dan tingkat interaksi media sosial yang relatif rendah. Cluster penggunaan sedang memiliki centroid akhir sebesar (7,25; 3,25; 5,88) yang mencerminkan aktivitas media sosial dengan intensitas menengah, sedangkan cluster penggunaan tinggi memiliki centroid akhir sebesar (12,25; 5,75; 8,88) yang menunjukkan frekuensi akses tinggi, durasi penggunaan yang lebih lama, serta tingkat interaksi media sosial yang tinggi. Hasil ini menunjukkan bahwa algoritma K-Means mampu mengelompokkan perilaku penggunaan media sosial siswa SMA secara efektif berdasarkan perbedaan karakteristik numerik pada setiap indikator yang digunakan.

#### REFERENCES

- [1] M. Hammam and T. Utomo, "CLUSTERING DATA SISWA PUTUS SEKOLAH DENGAN ALGORITMA K-MEANS DAN DBSCAN," *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 18, no. 2, pp. 150–159, 2023.
- [2] M. Sosial and D. Mining, "Pengelompokan ( Clustering ) Perilaku Penggunaan Media Sosial Siswa SMA Menggunakan Algoritma K-Means," *Inform. J. Ilmu Komput.*, vol. 3, no. 1, pp. 300–310, 2021.
- [3] E. W. Ambarsari, D. Fathudin, and G. Alfiani, "Utilizing K-Means Clustering to Understanding Audience Interest in SEO-Optimized Media Content," *J. Comput. Informatics Res.*, vol. 3, no. 2, pp. 208–214, 2024, doi: 10.47065/comforch.v3i2.1207.
- [4] S. Dewi, A. Kresnawati, S. Sopyanti, and A. Sulmainah, "Mapping Library User Behavior Base On K-Means Clustering Of Ma ' soem University Student Pemetaan Perilaku Pengguna Perpustakaan Berbasis K-Means Pada Mahasiswa Universitas Ma ' soem," *Bimaster Bul. Ilm. Mat. Stat. dan Ter.*, vol. 5, no. 2, pp. 130–136, 2025.
- [5] A. N. Fauzan, F. A. Wandu, A. Zuhair, N. Aiman, and M. Wati, "Pengelompokan Minat Akademik Siswa SMA Negeri 1 Loa Janan Menggunakan Metode Clustering K-means," *J. Bina Komput.*, vol. 9, no. 2, pp. 165–174, 2025.
- [6] A. E. Widjaja, A. Fransisko, and C. A. Haryani, "Text Mining Application with K-Means Clustering to Identify Sentiments and Popular Topics : a Case Study of the three Largest Online Marketplaces in Indonesia," *Inform. J. Ilmu Komput.*, vol. 4, no. 4, pp. 441–453, 2023.
- [7] Y. Malode, A. Aylani, A. Bhardwaj, and D. Hajoary, "Comparative Analysis of Community Detection Algorithms on the SNAP Social Circles Dataset," *Int. J. Environ. Res. Public Health*, vol. 1, no. 1, pp. 1–20, 2022.
- [8] D. M. Putra and F. F. Abdulloh, "Comparison of Clustering Algorithms : Fuzzy C-Means , K-Means , and DBSCAN for House Classification Based on Specifications and Price," *J. Appl. Informatics Comput.*, vol. 8, no. 2, pp. 509–515, 2024.
- [9] M. N. Afrilia, N. Rahaningsih, R. D. Dana, and N. D. Nuris, "Optimasi Analisis Clustering Untuk Aktivitas Dan Respon Pengguna Media Sosial Dengan K-Means," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 148–155, 2024.
- [10] P. K. Berindikasi, M. A. Syawal, and B. A. Pradhana, "Implementation Of The K-Means Method For Grouping Words Indicative Of Cyberbullying In," *Inform. J. Tek. Inform. dan Multimed.*, vol. 9, no. 2, pp. 330–347, 2025, doi: 10.52362/jisicom.v9i2.2163.
- [11] R. Dila, S. Defit, and S. Arlis, "Analisis Algoritma K-Means Clustering Dalam Pengelompokan Prestasi Belajar Siswa Menengah Atas ( SMA )," vol. 5, no. 5, pp. 1113–1119, 2025, doi: 10.47065/bulletincsr.v5i5.751.
- [12] F. Raihan, "Analisis Pengaruh Media Sosial X Terhadap Sikap dan Perilaku Remaja Menggunakan K-Means Clustering," *J. Indones. Manaj. Inform. Dan Komunikasi*, vol. 5, no. 3, pp. 3138–3146, 2024.
- [13] N. U. NLPP Dewi, IN Purnama, "Klasterisasi Mahasiswa Berdasarkan Performa Akademik Menggunakan Algoritma K-Means pada RapidMiner," *J. Eng. Informatics*, vol. 1, no. 1, pp. 46–50, 2025.
- [14] J. Evans, Y. Nataliani, S. Informasi, U. Kristen, S. Wacana, and J. O. N. Salatiga, "Analisis Perilaku Siswa pada Pembelajaran Daring Menggunakan k-Means Clustering dari Log Data Moodle pada SMPN 6 Salatiga," *JuSiTik J. Sist. dan Teknol. Inf. Komun.*, vol. 7, no. 2, pp. 43–49, 2024.
- [15] F. Aziz, B. K. Romadhan, and I. Z. Pratama, "Pengelompokan Berita Berdasarkan Kategori Like Action Pada Media Sosial

- Instagram Penamuda Menggunakan Algoritma K- means,” *JuSiTik J. Sist. dan Teknol. Inf. Komun.*, vol. 1, no. 2, pp. 57–60, 2024.
- [16] C. Agusniar, S. M. Ulfa, and H. Rhomadhona, “Utilizing K-Means Clustering for Grouping Student Achievement Data to Evaluate Learning Activeness,” *J. Adv. Comput. Knowl. Algorithms*, vol. 2, no. 3, pp. 56–59, 2025.
- [17] M. Risnasari, N. Aulia, and L. Cahyani, “Clustering Of Student Learning Styles in the Industry 4 . 0 Using KMeans Algorithm,” *J. Bina Komput.*, vol. 24, no. August, pp. 246–257, 2022.
- [18] Z. Gong and Y. Guo, “Social media use and academic performance among college students : the chain mediating roles of social anxiety and fear of missing out and the moderating effect of teacher-student relationship,” *J. Armada Inform.*, no. October, pp. 1–11, 2025, doi: 10.3389/fpsyg.2025.1649890.
- [19] O. Rizan, “The Influence of Social Media on Student Learning Behavior and Its Effects on Academic Achievement,” *Inform. J. Tek. Inform. dan Multimed.*, vol. 14, pp. 244–249, 2025.
- [20] B. Long, “Social Media and Students ’ Academic Performance : A Case Study of Students Pursuing Business Administration at Higher Education Institution in Phnom Penh , Cambodia,” *J. Armada Inform.*, vol. 2, 2023.