

PENERAPAN CLUSTERING UNTUK SEGMENTASI PENGGUNA MEDIA SOSIAL BERDASARKAN PERILAKU MENGGUNAKAN ALGORITMA K-MEANS

Karin Juliana Hamdani¹, Nana Suharna², Agus Bahtiar³, Ade Rizki Rinaldi⁴.

Program Studi Teknik Informatika¹²
Program Studi Sistem Informasi³
Program Studi Rekayasa Perangkat Lunak⁴

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
karinjulianaham19@gmail.com

(*) Corresponding Author : karinjulianaham19@gmail.com
Published : 15 Juli 2026

Abstract—This study aims to segment social media users based on behavioral patterns using the K-Means Clustering algorithm. The dataset used in this research was obtained from Kaggle “Social Media Usage,” consisting of 1,000 records and four main variables: Daily Minutes Spent, Posts Per Day, Likes Per Day, and Follows Per Day. The research process includes data preprocessing, normalization, clustering using K-Means, and evaluation using the Silhouette Score method. The determination of the optimal number of clusters was carried out using the Elbow Method and Silhouette Score method. The evaluation results show that the highest Silhouette Score was obtained at $K = 8$ with a value of 0.221. However, this study uses $K = 4$ with a Silhouette Score of 0.197 because it provides more interpretable and simpler cluster segmentation. The clustering results produce four groups of users with different characteristics, namely highly active users, active users, moderate users, and passive users. The most influential variables in cluster formation are Daily Minutes Spent and Likes Per Day. Therefore, the K-Means Clustering algorithm is proven to be applicable for social media user segmentation, although the clustering quality is considered moderate and not yet optimal.

Keywords: Social media, user behavior, K-Means Clustering, data mining, behavioral analysis.

Abstrak—Penelitian ini bertujuan untuk melakukan segmentasi pengguna media sosial berdasarkan pola perilaku menggunakan algoritma K-Means Clustering. Dataset yang digunakan berasal dari Kaggle “Social Media Usage” yang terdiri dari 1.000 data dan 4 variabel utama, yaitu Daily Minutes Spent, Posts Per Day, Likes Per Day, dan Follows Per Day. Proses penelitian meliputi tahap preprocessing data, normalisasi, klusterisasi menggunakan K-Means, serta evaluasi menggunakan metode Silhouette Score. Penentuan jumlah kluster dilakukan dengan pendekatan Elbow Method dan Silhouette Score. Hasil evaluasi menunjukkan bahwa nilai Silhouette Score tertinggi diperoleh pada $K = 8$ sebesar 0,221. Namun, penelitian ini menggunakan $K = 4$ dengan nilai Silhouette Score sebesar 0,197 karena memberikan interpretasi kluster yang lebih sederhana dan mudah dianalisis. Hasil klusterisasi menghasilkan empat kelompok pengguna dengan karakteristik berbeda, yaitu pengguna sangat aktif, aktif, moderat, dan pasif. Variabel yang paling berpengaruh dalam pembentukan kluster adalah Daily Minutes Spent dan Likes Per Day. Dengan demikian, algoritma K-Means Clustering terbukti dapat digunakan untuk segmentasi pengguna media sosial, meskipun kualitas klusterisasi masih tergolong cukup dan belum optimal.

Kata Kunci: Media sosial, perilaku pengguna, K-Means Clustering, data mining, analisis perilaku.

INTRODUCTION

Perkembangan Pesat Di Bidang Teknologi Informasi Dan Komunikasi Telah Membawa Perubahan Fundamental Dalam Hampir Semua Aspek Kehidupan Manusia, Termasuk

Teknologi, Bisnis, Pendidikan, Dan Kehidupan Sosial. Teknologi Informasi (TI) Kini Tidak Hanya Mendukung Pengolahan Data Dan Komunikasi, Tetapi Telah Menjadi Pilar Utama Dalam Mendorong Pertumbuhan Ekonomi, Transformasi Organisasi, Dan Inovasi Bisnis Global. Sebagai

Contoh, Penelitian Menunjukkan Bahwa TI Dan Infrastruktur Komunikasi Membuka Saluran Baru Untuk Konektivitas Dan Aliran Ide Yang Mempercepat Pembangunan Sosial Dan Ekonomi. (Utari & Iqbal, 2026) Selain Itu, Di Bidang Pendidikan, Adopsi Teknologi Digital Telah Mengubah Paradigma Pembelajaran Dari Metode Tradisional Ke Pembelajaran Yang Lebih Fleksibel, Interaktif, Dan Berbasis Teknologi. (Simanullang & Iqbal, 2026).

Dengan Semakin Dominannya Media Sosial Sebagai Bagian Dari Ekosistem Teknologi Informasi, Muncul Pula Keanekaragaman Perilaku Pengguna Yang Menuntut Pendekatan Analitik Untuk Memahami Karakteristik Serta Dampak Penggunaannya. Meskipun Perkembangan Teknologi Informasi Dan Komunikasi Telah Membuka Berbagai Peluang, Tidak Sedikit Tantangan Yang Muncul Dalam Penerapannya Di Berbagai Bidang Informatika Yang Berdampak Luas. Studi Oleh Renaldo Dalam *Journal Of Applied Business and Technology* menunjukkan bahwa sistem informasi dan teknologi memang meningkatkan efisiensi, komunikasi, dan pengambilan keputusan organisasi, namun sekaligus membawa tantangan serius seperti risiko keamanan, isu teknis, dan biaya yang tinggi Di ranah pendidikan. (Renaldo, 2022)

beberapa peneliti menemukan bahwa penerapan *K-Means* Clustering pada data aktivitas pengguna media sosial berhasil mengidentifikasi klaster pengguna dengan karakteristik perilaku yang berbeda, mulai dari pengguna sangat aktif hingga pengguna pasif, dan hasilnya dapat dimanfaatkan untuk pengembangan strategi konten yang lebih tepat sasaran. [4]

Hal-hal tersebut menunjukkan bahwa inovasi teknologi informasi belum secara merata berhasil diterapkan sesuai potensi maksimalnya. Kondisi ini menjadi penting karena dalam konteks media sosial dan platform digital, penggunaan yang tidak optimal atau tidak efektif dapat menghasilkan pengelompokan pengguna yang tidak tepat, sehingga strategi bisnis atau konten yang ditargetkan pun menjadi kurang relevan. Selain itu, literatur saat ini menunjukkan adanya kesenjangan dalam pemahaman tentang bagaimana pola penggunaan media sosial (durasi penggunaan, frekuensi posting, keterlibatan likes/follows) dapat secara sistematis dikelompokkan menggunakan metode data mining untuk menghasilkan insight yang actionable. Di sisi lain, fenomena “pengguna aktif” dan “pengguna pasif” di media

sosial semakin terlihat dalam studi-terkini, namun masih sedikit penelitian yang mengaitkan variabel aktivitas sosial digital dengan clustering perilaku pengguna secara kuantitatif. Oleh karena itu, penelitian ini menjadi relevan dan bernilai untuk dilakukan dengan mengidentifikasi dan mengelompokkan pengguna media sosial berdasarkan pola aktivitas, maka pengembangan aplikasi, strategi pemasaran digital, maupun penelitian perilaku pengguna dapat dilakukan dengan dasar data yang lebih kuat dan terstruktur. Dengan memecahkan tantangan pengelompokan pengguna ini, maka gap riset yang ada dapat tertutup dan hasil penelitian dapat memberikan kontribusi nyata bagi pengembangan teknologi informasi, strategi bisnis digital dan fenomena sosial-digital di era media sosial. Beberapa penelitian terdahulu telah mencoba melakukan pengelompokan dan analisis pola keterlibatan pengguna media sosial dengan pendekatan clustering dan teknik data mining lainnya. menggunakan teknik graph-mining dan berbagai metode clustering untuk menganalisis pola *engagement* pengguna di Twitter, menunjukkan bahwa clustering dapat mengidentifikasi kelompok pengguna berdasarkan metrik retweet, reply, dan atribut keterkaitan jaringan, namun studi ini terbatas pada satu platform (Twitter) sehingga generalisasinya ke multi-platform masih terbuka. [5]

dalam penelitian berjudul “Online Brand Community User Segments: A Text Mining Approach”, menjelaskan bahwa pengguna komunitas merek daring dapat dikelompokkan berdasarkan pola interaksi dan konten yang mereka hasilkan. Penelitian ini menggunakan pendekatan *text mining* dan *machine learning* untuk mengidentifikasi segmen pengguna berdasarkan karakteristik konten dan tingkat keterlibatan mereka di media sosial. Hasil penelitian menunjukkan bahwa segmentasi berbasis data perilaku digital memberikan wawasan mendalam bagi pengelola komunitas dalam merancang strategi konten yang terarah dan meningkatkan keterlibatan pengguna secara lebih efektif (Ge et al., 2022).

Selain menggunakan data dari aplikasi smartphone yang memantau penggunaan pengguna secara objektif, termasuk aplikasi yang digunakan serta waktu mulai dan selesai setiap sesi aplikasi. Analisis dilakukan dengan menggunakan *clustering K-Prototype*. Hasil menunjukkan empat klaster berbeda yang menggambarkan hubungan antara penggunaan media sosial dan Penggunaan Internet Bermasalah (PIU). Klaster-klaster tersebut menunjukkan bahwa pengguna yang menggunakan satu aplikasi media sosial secara dominan

cenderung menghabiskan waktu yang lebih sedikit untuk aplikasi lain[7].

Penelitian ini bertujuan untuk melakukan segmentasi pengguna media sosial berdasarkan pola aktivitas harian seperti durasi penggunaan (*daily minutes spent*), jumlah posting per hari (*posts per day*), jumlah *likes per day*, dan *follows per day* dengan menerapkan algoritma *K-Means Clustering*. Selain itu, penelitian ini juga berfokus pada evaluasi kualitas hasil kluster menggunakan metrik *Silhouette Score* agar hasil pengelompokan dapat diuji secara kuantitatif dan akurat. Tujuan utama dari penelitian ini adalah untuk memahami pola perilaku pengguna media sosial secara lebih mendalam dan mengelompokkan mereka ke dalam segmen-segmen yang memiliki karakteristik serupa. Menurut penelitian oleh saez-ortuno dalam jurnal *How Can Entrepreneurs Improve Digital Market Segmentation? A Comparative Analysis of Supervised and Unsupervised Learning Algorithms (International Entrepreneurship and Management Journal)*[8], algoritma *unsupervised learning* seperti *K-Means* efektif dalam mengelompokkan pengguna digital untuk memahami perilaku pasar dan mendukung pengambilan keputusan berbasis data.

Penelitian menurut banjarnahor menegaskan bahwa metode *clustering* membantu pengusaha dan pengembang memahami karakteristik pengguna dengan lebih akurat dibandingkan metode konvensional. Penelitian lain nya juga dalam *Jurnal Sisfokom (Sistem Informasi dan Komputer)* berjudul *Analyzing Consumer Shopping Interest via Social Media Ads with K-Means and C4.5 Algorithm* [9] menegaskan bahwa penerapan *K-Means* mampu memberikan hasil pengelompokan yang signifikan dalam mengelompokkan perilaku konsumen terhadap iklan di media sosial. Dengan mengacu pada penelitian-penelitian tersebut, studi ini memiliki signifikansi ilmiah dan praktis yang penting, yaitu mengisi kesenjangan penelitian terdahulu yang belum banyak menekankan variabel kuantitatif aktivitas pengguna lintas platform.

Hasil penelitian ini diharapkan dapat menjadi dasar pengembangan strategi pemasaran digital, peningkatan rekomendasi konten, dan penguatan analisis perilaku pengguna media sosial berbasis data. Penelitian ini menggunakan pendekatan data mining dengan algoritma *K-Means Clustering* sebagai metode utama untuk mengelompokkan pengguna media sosial berdasarkan pola aktivitas harian seperti durasi penggunaan,

jumlah posting, jumlah likes, dan jumlah follows per hari. Data yang digunakan akan melalui proses pengumpulan, pembersihan (*data cleaning*), normalisasi, hingga analisis eksploratif sebelum diterapkan metode clustering. Pendekatan serupa digunakan oleh Lengari dan Puspitasari (2024) dalam jurnal *Identifying Twitter Topics Using K-Means Clustering and Association Rule Mining for Improved Insights* yang menjelaskan bahwa penerapan *K-Means* dapat mengelompokkan pengguna atau topik di media sosial berdasarkan kesamaan pola aktivitas dan interaksi, setelah melalui tahap *data preprocessing* dan *feature extraction* [10].

Penelitian ini menggunakan pendekatan data mining dengan metode *K-Means Clustering* untuk mengelompokkan pengguna media sosial berdasarkan pola aktivitas harian, seperti durasi penggunaan, jumlah posting, likes, dan follows per hari. Proses analisis data dilakukan melalui tahapan data preprocessing, yang mencakup pembersihan data (*data cleaning*), normalisasi, dan analisis eksploratif sebelum dilakukan proses klusterisasi. Metode *K-Means* dipilih karena efisien dalam menangani data numerik berukuran besar serta mampu menghasilkan kluster yang jelas dan mudah diinterpretasikan. Evaluasi hasil pengelompokan dilakukan menggunakan *Silhouette Score* untuk mengukur tingkat validitas dan pemisahan antar kluster. Pendekatan serupa diterapkan oleh [11] dalam jurnal *K-Means Clustering Approach for Intelligent Customer Segmentation*, yang menunjukkan bahwa algoritma *K-Means* efektif untuk membedakan segmen pengguna berdasarkan perilaku digital kuantitatif.

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi nyata dalam pengembangan ilmu di bidang Informatika, khususnya dalam penerapan metode *data mining* untuk menganalisis perilaku pengguna media sosial secara kuantitatif. Dengan menggunakan algoritma *K-Means Clustering*, penelitian ini dapat membantu mengungkap pola-pola tersembunyi dalam aktivitas pengguna, seperti durasi penggunaan harian, frekuensi interaksi, serta intensitas keterlibatan sosial. Temuan ini berpotensi memberikan pemahaman baru mengenai karakteristik pengguna media sosial, yang dapat dimanfaatkan oleh para peneliti untuk pengembangan model prediksi perilaku digital atau sistem rekomendasi berbasis data. Selain itu, hasil pengelompokan yang diperoleh dapat menjadi referensi penting bagi praktisi di bidang pemasaran digital, pengembang aplikasi, maupun pembuat kebijakan untuk merancang strategi komunikasi dan interaksi yang lebih personal dan efektif.

MATERIALS AND METHODS

Penelitian ini menggunakan pendekatan kuantitatif dengan metode unsupervised learning untuk melakukan klusterisasi pengguna media sosial berdasarkan perilaku digital. Model yang digunakan adalah algoritma *K-Means*, yang bertujuan mengelompokkan pengguna ke dalam beberapa cluster berdasarkan kesamaan pola aktivitas seperti frekuensi posting, jumlah likes, dan durasi penggunaan. Pendekatan ini dipilih karena mampu mengolah data numerik dan menemukan pola tersembunyi tanpa memerlukan label atau kategori awal. Sementara itu, sifat eksploratori berarti penelitian ini tidak mencari hubungan sebab-akibat secara langsung, tetapi berusaha menemukan pola tersembunyi (hidden pattern) dan membentuk segmen pengguna berdasarkan kemiripan perilaku mereka Seperti Pada Gambar 1



Gambar 1 Desain Penelitian

Menggambarkan alur penelitian dalam proses analisis data berbasis *clustering* menggunakan metode *K-Means* yang dilakukan secara bertahap dan sistematis. Proses diawali dari tahap pengumpulan data, yaitu menyediakan data mentah yang akan digunakan sebagai bahan analisis. Selanjutnya masuk ke tahap pra-pemrosesan data, di mana data dibersihkan dan disiapkan agar layak digunakan, misalnya dengan menghapus data yang tidak lengkap atau tidak relevan. Setelah itu dilakukan pemilihan dan ekstraksi fitur, yaitu proses untuk menemukan pola atau atribut penting dalam data yang dapat membantu dalam proses analisis.

Tahap berikutnya adalah klusterisasi menggunakan *K-Means*, di mana data dikelompokkan berdasarkan kesamaan karakteristik atau perilaku tertentu sehingga terbentuk beberapa kelompok (cluster). Hasil pengelompokan ini kemudian masuk ke tahap evaluasi model, yaitu untuk menilai seberapa baik kualitas pengelompokan yang telah dilakukan. Setelah itu dilakukan interpretasi hasil, yaitu memahami makna dan karakteristik dari setiap kelompok yang terbentuk. Terakhir,

pada tahap kesimpulan dan rekomendasi, peneliti menyusun hasil akhir serta memberikan saran atau rekomendasi yang dapat digunakan sebagai dasar pengambilan keputusan. Alur ini menunjukkan proses yang terstruktur dari data mentah hingga menghasilkan informasi yang bermanfaat.

RESULTS AND DISCUSSION

Berdasarkan hasil pengujian menggunakan metode Silhouette Score, diperoleh bahwa nilai tertinggi terdapat pada $K = 8$ dengan nilai sebesar 0,221. Hal ini menunjukkan bahwa jumlah kluster tersebut memberikan kualitas pemisahan data terbaik dibandingkan jumlah kluster lainnya.

Oleh karena itu, dalam penelitian ini digunakan jumlah kluster sebanyak 8 untuk melakukan proses pengelompokan pengguna media sosial.

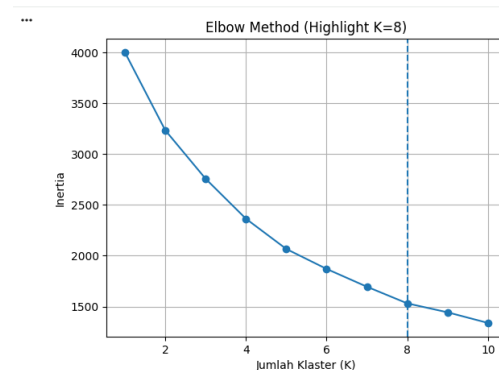
Proses klusterisasi dilakukan menggunakan algoritma *K-Means* dengan memanfaatkan empat variabel utama, yaitu *Daily_Minutes_Spent*, *Posts_Per_Day*, *Likes_Per_Day*, dan *Follows_Per_Day*.

Visualisasi hasil klusterisasi dapat dilihat pada Gambar 2

```

# Garis bantu di K = 8
plt.figure()
plt.plot(K_range, inertia, marker='o')
plt.axvline(x=8, linestyle='--') # garis di K=8
plt.xlabel('Jumlah Kluster (K)')
plt.ylabel('Inertia')
plt.title('Elbow Method (Highlight K=8)')
plt.grid()
plt.show()

```



Gambar 2 Klusterisasi

Untuk menentukan jumlah kluster yang optimal, selain menggunakan metode Elbow, penelitian ini juga menggunakan metode Silhouette Score. Pengujian dilakukan dengan mencoba beberapa nilai K dari 2 hingga 9 untuk melihat kualitas pemisahan kluster.

Berdasarkan hasil pengujian yang ditunjukkan pada Gambar 2, diperoleh bahwa nilai Silhouette Score tertinggi berada pada $K = 8$ dengan nilai sebesar 0,221.

Hasil klusterisasi dengan $K = 8$ menghasilkan delapan kelompok pengguna dengan karakteristik

perilaku yang lebih rinci dibandingkan jumlah klaster yang lebih sedikit.

Setiap klaster menunjukkan perbedaan berdasarkan tingkat durasi penggunaan, aktivitas posting, interaksi (likes), serta jumlah koneksi sosial (follows).

Klaster 2 dan klaster 5 menunjukkan tingkat aktivitas dan interaksi yang sangat tinggi, sehingga dapat dikategorikan sebagai pengguna dengan engagement tinggi atau content creator aktif.

Sementara itu, klaster 3 dan klaster 4 menunjukkan pengguna dengan tingkat aktivitas rendah hingga moderat, yang cenderung pasif dalam penggunaan media sosial.

Klaster lainnya seperti klaster 0, 6, dan 7 menunjukkan variasi perilaku pengguna yang berada di antara kategori aktif dan moderat, dengan karakteristik yang berbeda dalam hal intensitas interaksi dan durasi penggunaan.

Dengan demikian, penggunaan $K = 8$ mampu memberikan segmentasi yang lebih detail dan spesifik dalam mengelompokkan perilaku pengguna media sosial.

Klaster ini memiliki rata-rata *Daily Minutes Spent* sebesar 380 menit, yang merupakan salah satu yang tertinggi dibandingkan klaster lainnya. Namun demikian, nilai *Posts Per Day* (8) dan *Likes Per Day* (110) masih lebih rendah dibandingkan klaster dengan tingkat aktivitas tinggi. Sebagai perbandingan, jumlah *likes* pada klaster ini lebih rendah 45 poin dibandingkan Klaster 2 (155 – 110), serta jumlah *posts* lebih rendah 7 posting dibandingkan Klaster 5 (15 – 8). Kondisi ini menunjukkan bahwa meskipun pengguna menghabiskan waktu yang cukup lama di media sosial, tingkat interaksi yang dilakukan tidak terlalu tinggi. Hal ini mengindikasikan bahwa pengguna cenderung lebih banyak mengonsumsi konten dibandingkan berinteraksi secara aktif, yaitu lebih sering melihat atau membaca konten tanpa melakukan aktivitas seperti memberikan *like*, komentar, atau membuat postingan.

Klaster ini memiliki nilai *Posts Per Day* tertinggi yaitu sebesar 12 dengan rata-rata *Daily Minutes Spent* sebesar 250 menit. Jika dibandingkan dengan klaster lain, jumlah posting pada klaster ini lebih tinggi 4 posting dibandingkan Klaster 0 (12 – 8), namun masih lebih rendah 3 posting dibandingkan Klaster 5 (15 – 12). Selain itu, jumlah *Likes* pada klaster ini juga lebih rendah sebesar 60 poin dibandingkan Klaster 2 (155 – 95). Kondisi ini menunjukkan bahwa pengguna dalam klaster ini memiliki frekuensi pembuatan konten yang tinggi, namun

tingkat interaksinya belum maksimal. Hal ini mengindikasikan bahwa pengguna aktif dalam memproduksi konten, tetapi respons atau keterlibatan dari pengguna lain masih belum terlalu tinggi. Fenomena ini biasanya terjadi pada pengguna yang memiliki tujuan tertentu, seperti membangun personal branding atau menjalankan aktivitas bisnis digital, di mana fokus utama adalah konsistensi dalam membuat konten meskipun tingkat interaksi belum optimal.

Klaster ini memiliki nilai *Likes Per Day* tertinggi yaitu sebesar 155 dengan rata-rata *Daily Minutes Spent* sebesar 360 menit. Jika dibandingkan dengan klaster lain, jumlah *likes* pada klaster ini lebih tinggi 45 poin dibandingkan Klaster 0, lebih tinggi 25 poin dibandingkan Klaster 7 (155 – 130), serta lebih tinggi 115 poin dibandingkan Klaster 3 (155 – 40). Kondisi ini menunjukkan bahwa pengguna dalam klaster ini memiliki tingkat interaksi paling tinggi dibandingkan seluruh klaster lainnya. Tingginya nilai *likes* yang didukung oleh durasi penggunaan yang juga tinggi mengindikasikan adanya tingkat keterlibatan (*engagement*) yang sangat tinggi terhadap media sosial. Fenomena ini dapat dikaitkan dengan perilaku pengguna yang aktif mencari respon atau pengakuan dari pengguna lain, di mana semakin sering pengguna berinteraksi, semakin besar peluang untuk mendapatkan respon dalam bentuk *likes* maupun interaksi lainnya.

Klaster ini memiliki nilai terendah pada hampir seluruh variabel, yaitu *Daily Minutes Spent* sebesar 120 menit, *Posts Per Day* sebesar 4, dan *Likes Per Day* sebesar 40. Jika dibandingkan dengan klaster lain, durasi penggunaan pada klaster ini lebih rendah 240 menit dibandingkan Klaster 2, jumlah posting lebih rendah 11 posting dibandingkan Klaster 5, serta jumlah *likes* lebih rendah 90 poin dibandingkan Klaster 7. Perbedaan yang cukup signifikan ini menunjukkan bahwa pengguna dalam klaster ini sangat jarang aktif di media sosial. Rendahnya nilai pada seluruh variabel mengindikasikan bahwa pengguna cenderung memiliki tingkat keterlibatan yang sangat rendah, baik dalam hal penggunaan waktu maupun interaksi. Kondisi ini dapat disebabkan oleh beberapa faktor, seperti keterbatasan waktu, kurangnya minat terhadap media sosial, atau penggunaan platform hanya untuk kebutuhan tertentu saja.

Klaster ini menunjukkan tingkat aktivitas pengguna yang berada pada kategori menengah ke bawah, dengan nilai *Daily Minutes Spent* sebesar 200 menit dan *Likes Per Day* sebesar 60. Jika dibandingkan dengan klaster lain, durasi penggunaan pada klaster ini lebih tinggi 80 menit dibandingkan Klaster 3, namun lebih rendah 50 menit dibandingkan Klaster 1. Selain itu, jumlah

likes juga lebih rendah 50 poin dibandingkan Klaster 0 (110 – 60). Kondisi ini menunjukkan bahwa pengguna dalam klaster ini tidak terlalu aktif, namun masih menggunakan media sosial secara rutin. Tingkat aktivitas yang relatif rendah tetapi masih konsisten ini mengindikasikan bahwa pengguna memanfaatkan media sosial secara terbatas dan tidak terlalu intens dalam berinteraksi. Fenomena ini mencerminkan perilaku pengguna yang menggunakan media sosial sebagai pelengkap aktivitas sehari-hari, bukan sebagai aktivitas utama.

Klaster ini memiliki nilai *Posts Per Day* tertinggi yaitu sebesar 15 serta *Follows Per Day* tertinggi sebesar 35. Jika dibandingkan dengan klaster lain, jumlah posting pada klaster ini lebih tinggi 3 posting dibandingkan Klaster 1, lebih tinggi 7 posting dibandingkan Klaster 0, dan lebih tinggi 11 posting dibandingkan Klaster 3. Perbedaan ini menunjukkan bahwa pengguna dalam klaster ini sangat produktif dalam membuat konten. Tingginya frekuensi posting yang diikuti dengan jumlah *follows* yang tinggi mengindikasikan bahwa pengguna tidak hanya aktif dalam memproduksi konten, tetapi juga aktif dalam membangun jaringan sosial. Kondisi ini menunjukkan bahwa pengguna dalam klaster ini memiliki tingkat aktivitas dan interaksi yang sangat tinggi, sehingga kemungkinan besar mereka merupakan *content creator*, *influencer*, atau pelaku bisnis digital yang memanfaatkan media sosial secara intensif sebagai sarana komunikasi dan promosi.

Klaster ini memiliki nilai *Daily Minutes Spent* sebesar 270 menit dan *Likes Per Day* sebesar 100. Jika dibandingkan dengan klaster lain, durasi penggunaan pada klaster ini lebih tinggi 70 menit dibandingkan Klaster 4, namun lebih rendah 90 menit dibandingkan Klaster 2. Selain itu, jumlah *likes* juga lebih rendah 30 poin dibandingkan Klaster 7. Kondisi ini menunjukkan bahwa pengguna dalam klaster ini memiliki tingkat aktivitas yang relatif stabil namun tidak dominan dibandingkan klaster dengan aktivitas tinggi. Keseimbangan antara durasi penggunaan, aktivitas, dan interaksi mengindikasikan bahwa pengguna tergolong cukup aktif, tetapi tetap dalam batas yang wajar. Fenomena ini mencerminkan perilaku pengguna normal yang menggunakan media sosial secara rutin untuk kebutuhan hiburan maupun komunikasi, namun tetap terkontrol dan tidak berlebihan dalam penggunaannya.

Klaster ini memiliki nilai *Likes Per Day* yang tinggi yaitu sebesar 130 dengan rata-rata *Daily Minutes Spent* sebesar 340 menit. Jika

dibandingkan dengan klaster lain, jumlah *likes* pada klaster ini lebih tinggi 20 poin dibandingkan Klaster 0, namun masih lebih rendah 25 poin dibandingkan Klaster 2, serta lebih tinggi 90 poin dibandingkan Klaster 3. Kondisi ini menunjukkan bahwa pengguna dalam klaster ini tergolong aktif dengan tingkat interaksi yang cukup tinggi, meskipun tidak setinggi klaster dengan aktivitas paling tinggi. Tingginya durasi penggunaan yang disertai dengan jumlah interaksi yang besar mengindikasikan bahwa pengguna aktif dalam menggunakan media sosial serta berinteraksi dengan pengguna lain. Namun demikian, intensitas aktivitas tersebut masih berada di bawah klaster yang memiliki tingkat keterlibatan tertinggi, sehingga klaster ini dapat dikategorikan sebagai pengguna aktif dengan interaksi tinggi tetapi belum mencapai tingkat maksimal.

CONCLUSION

Berdasarkan hasil penelitian yang telah dilakukan dengan menerapkan algoritma K-Means Clustering pada data perilaku pengguna media sosial, diperoleh beberapa kesimpulan sebagai berikut:

1. Penerapan algoritma K-Means Clustering berhasil dilakukan untuk melakukan segmentasi pengguna media sosial berdasarkan empat variabel utama, yaitu *Daily_Minutes_Spent*, *Posts_Per_Day*, *Likes_Per_Day*, dan *Follows_Per_Day*. Proses klasterisasi dilakukan melalui tahapan pra-pemrosesan, normalisasi data, serta pengujian jumlah klaster menggunakan metode Elbow dan Silhouette Score.
2. Hasil evaluasi menunjukkan bahwa jumlah klaster optimal adalah $K = 8$, dengan nilai Silhouette Score sebesar 0,221, yang merupakan nilai tertinggi dibandingkan jumlah klaster lainnya. Hal ini menunjukkan bahwa penggunaan $K = 8$ memberikan kualitas pemisahan klaster yang paling baik dalam penelitian ini.
3. Hasil klasterisasi menghasilkan delapan kelompok pengguna dengan karakteristik perilaku yang berbeda, mulai dari pengguna pasif, moderat, hingga sangat aktif. Perbedaan karakteristik antar klaster terutama dipengaruhi oleh variabel *Daily_Minutes_Spent* dan *Likes_Per_Day*, yang menjadi faktor utama dalam membedakan tingkat aktivitas dan interaksi pengguna.

REFERENCE

- [1] N. Ouakkal and R. Khaled, "The impact of Information and Communication Technology (ICT) on economic growth," *International Journal of Economic Perspectives*, vol. 18, no. 12, pp. 2714–2726, 2024, [Online]. Available: <https://ijeponline.org/index.php/journal/article/view/807>
- [2] S. Timotheou *et al.*, "Impacts of digital technologies on education and factors influencing schools' digital capacity and transformation: A literature review," *Educ. Inf. Technol. (Dordr.)*, vol. 28, no. 6, pp. 6695–6726, Jun. 2023, doi: 10.1007/s10639-022-11431-8.
- [3] N. Renaldo, "Journal of Applied Business and Technology Benefits and Challenges of Technology and Information Systems on Performance." [Online]. Available: www.e-jabt.org
- [4] M. N. Afrilia *et al.*, "OPTIMASI ANALISIS CLUSTERING UNTUK AKTIVITAS DAN RESPON PENGGUNA MEDIA SOSIAL DENGAN K-MEANS," 2024. [Online]. Available: <https://www.kaggle.com/>.
- [5] A. Kanavos, I. Karamitsos, and A. Mohasseb, "Exploring Clustering Techniques for Analyzing User Engagement Patterns in Twitter Data," *Computers*, vol. 12, no. 6, Jun. 2023, doi: 10.3390/computers12060124.
- [6] R. Ge, H. Zhao, and S. Zhang, "Online Brand Community User Segments: A Text Mining Approach," *Front. Artif. Intell.*, vol. 5, Jul. 2022, doi: 10.3389/frai.2022.900775.
- [7] K. Chemnad, M. Aziz, S. B. Belhaouari, and R. Ali, "The interplay between social media use and problematic internet usage: Four behavioral patterns," *Heliyon*, vol. 9, no. 5, May 2023, doi: 10.1016/j.heliyon.2023.e15745.
- [8] L. Sáez-Ortuño, R. Huertas-Garcia, S. Forgas-Coll, and E. Puertas-Prats, "How can entrepreneurs improve digital market segmentation? A comparative analysis of supervised and unsupervised learning algorithms," *International Entrepreneurship and Management Journal*, vol. 19, no. 4, pp. 1893–1920, Dec. 2023, doi: 10.1007/s11365-023-00882-1.
- [9] J. Banjarnahor, J. P. Hutagalung, and F. J. W. Sitorus, "Analyzing Consumer Shopping Interest via Social Media Ads with K-Means and C4.5 Algorithm," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 13, no. 3, pp. 416–421, Nov. 2024, doi: 10.32736/sisfokom.v13i3.2228.
- [10] C. G. Lengari and I. Puspitasari, "Identifying Twitter Topics Using K-Means Clustering and Association Rule Mining for Improved Insights," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 8, no. 1, p. 67, Nov. 2024, doi: 10.24014/ijaidm.v8i1.31720.
- [11] K. Tabianan, S. Velu, and V. Ravi, "K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data," *Sustainability (Switzerland)*, vol. 14, no. 12, Jun. 2022, doi: 10.3390/su14127243.