



Analisis Komparatif Karakteristik Konten Antar Kluster Engagement Menggunakan K-Means Clustering

Febriyola Desy Azzahra¹, Fitri Yunita², Bayu Rianto³

^{1,2,3} Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Islam Indragiri

febriyolada@gmail.com, fitriyun@gmail.com*, rianto.bayu91@gmail.com *

Abstrak

Penerapan teknologi digital untuk pemasaran Usaha Mikro, Kecil, dan Menengah (UMKM) sangat krusial bagi keberlanjutan usaha, sebagaimana ditegaskan dalam penelitian Kedi yang menyoroti pentingnya adopsi teknologi meskipun di tengah berbagai keterbatasan. Namun, evaluasi kinerja pemasaran digital pada UMKM lokal sering kali terkendala ketiadaan instrumen terukur berbasis data log, sehingga strategi konten masih bergantung pada asumsi konvensional. Penelitian ini bertujuan mengintegrasikan data log historis ke dalam sistem evaluasi terstruktur guna merumuskan strategi konten yang efektif bagi UMKM X di sektor F&B. Sebanyak 180 data unggahan publik bersih diekstrak dan dikelompokkan menggunakan algoritma *K-Means Clustering* yang dipadukan dengan kriteria kerangka kerja AIDA (Attention, Interest, Desire, Action). Validitas pemodelan diuji menggunakan *Silhouette Coefficient* berbasis *Engagement Rate* serta uji statistik ANOVA dan Chi-Square. Hasil klusterisasi berhasil membentuk tiga kelompok utama: kluster postingan fitur AIDA lengkap, kluster *engagement* tinggi dengan minim fitur AIDA, dan kluster menengah dengan *engagement* sedang hingga rendah. Temuan ini mengindikasikan bahwa format visual dinamis memicu interaksi tertinggi. Sebagai bentuk penerapan praktis sistem informasi, *centroid* hasil klusterisasi ditransformasikan ke dalam prototipe lembar kerja digital (*spreadsheet*) interaktif. Instrumen ini memberikan alat bantu evaluasi mandiri yang praktis dan berbiaya rendah bagi pemilik UMKM X untuk memprediksi serta mengoptimalkan performa unggahan baru secara terukur dan berbasis data.

Kata Kunci: Sistem Informasi Evaluasi, Analitik Data Log, K-Means Clustering, Model AIDA, Engagement Rate, UMKM

1. Pendahuluan

Pada era digital, sektor Food and Beverage (F&B) sangat bergantung pada visualisasi digital untuk menarik minat konsumen. Platform Instagram menjadi media utama pelaku usaha kuliner untuk menekan biaya pemasaran (marketing cost) sekaligus menjangkau target pasar secara presisi (Nasution, 2021). Melalui media sosial, perusahaan dapat membangun identitas produk dan menciptakan keterlibatan pelanggan (customer engagement) yang mendalam. Namun, fenomena di lapangan menunjukkan banyak pelaku usaha cenderung bertumpu pada strategi hard-selling yang agresif melalui format statis maupun carousel. Pendekatan ini sering kali berbanding terbalik dengan tingkat interaksi audiens, sedangkan pendekatan soft-selling justru lebih efektif dalam membangun emotional engagement (Salsabila, 2025). Mengacu pada model AIDA (*Attention, Interest, Desire, Action*), keberhasilan komunikasi digital memerlukan keseimbangan antara daya tarik visual awal untuk memicu *Attention* dan penyampaian pesan emosional/interaktif guna mendorong tahapan *Action* (Yanti, 2023).

Masalah tersebut teramati secara nyata pada akun Instagram UMKM X, sebuah pelaku usaha F&B lokal. Berdasarkan observasi awal terhadap sampel 20 unggahan terbaru, ditemukan 19 di antaranya merupakan konten berformat carousel dengan dominasi hard-selling yang memenuhi indikator Business Tag, Promo Language, Call to Action (CTA), dan Transaction Access. Selain itu, terdapat ketidakkonsistenan distribusi waktu unggah yang memuncak hanya pada periode tertentu. Dominasi promosi masif yang tidak teratur ini menghasilkan respons audiens yang rendah (rata-rata 45 likes), mengindikasikan adanya ketidakefisienan dalam strategi komunikasi dan manajemen konten. Untuk mengatasi permasalahan ini, evaluasi berbasis data objektif (*data-driven insights*) sangat diperlukan agar pengambilan keputusan tidak mengandalkan intuisi semata (Rahman et al., 2025). Pengolahan data log historis interaksi publik menggunakan teknik data mining dan kerangka kerja *Knowledge Discovery in Databases* (KDD) menawarkan solusi terstruktur untuk membedah karakteristik pesan dan pola performa konten (Lubis & Iqbal, 2026).

Beberapa penelitian terdahulu telah mengkaji strategi media sosial dan klusterisasi konten. Penelitian-penelitian sebelumnya membuktikan efektivitas algoritma *K-Means Clustering* dalam mengelompokkan data aktivitas media sosial (Ambarsari et al., 2024; Nafsyah et al., 2022), serta menekankan pentingnya pengujian *Silhouette Coefficient* untuk validasi kelompok dan penerapan kerangka AIDA (Nurhidayat et al., 2025; Yadav et al., 2025). Meskipun demikian, penelitian terdahulu umumnya menganalisis aspek komputasional (*data mining*) dan aspek teori komunikasi pemasaran secara terpisah. Belum ada studi yang secara komprehensif mengintegrasikan metode K-Means Clustering berbasis AIDA dengan validasi statistik ganda (*Silhouette Coefficient*, ANOVA, dan Chi-Square) untuk menghasilkan luaran aplikatif berupa instrumen evaluasi mandiri bagi UMKM F&B lokal. Hal ini

menimbulkan kesenjangan (research gap) antara ketersediaan data log publik dengan penerapannya secara praktis bagi efisiensi bisnis kecil (Ambarsari et al., 2024; Kedi et al., 2024)

Untuk menutup kesenjangan tersebut, penelitian ini bertujuan mengelompokkan dan mengevaluasi 180 data unggahan historis publik pada akun UMKM X secara komputasional mengadaptasi siklus KDD. Penelitian ini menerapkan algoritma K-Means Clustering berdasarkan metrik Engagement Rate dan fitur variabel AIDA [11], yang divalidasi menggunakan Silhouette Coefficient serta uji ANOVA dan Chi-Square [12]. Kebaruan (novelty) dari penelitian ini terletak pada pendekatan integratif (Double-Layer Analysis) yang mengombinasikan pemodelan komputasional K-Means dengan kerangka AIDA, serta penyusunan prototipe lembar kerja digital (spreadsheet). Luaran ini dirancang agar pemilik akun UMKM X dapat memprediksi dan mengevaluasi kinerja konten secara mandiri berbasis data terukur secara efisien (Arief et al., 2026).

2. Metode Penelitian

Penelitian ini menerapkan pendekatan metode campuran (*mixed methods*) berbasis kerangka kerja *Knowledge Discovery in Databases* (KDD) (Nallasivam et al., 2025) yang dijalankan secara sekuensial. Aspek kuantitatif diterapkan melalui pengelompokkan data log menggunakan algoritma *K-Means Clustering* yang divalidasi dengan *Silhouette Coefficient*, uji *Analysis of Variance* (ANOVA), dan uji *Chi-Square*. Aspek kualitatif diterapkan melalui analisis isi (*content analysis*) berbasis model hierarki komunikasi AIDA (*Attention, Interest, Desire, Action*) untuk merumuskan rekomendasi strategi (Yadav et al., 2025).

2.1. Bahan dan Pengumpulan Data

Subjek penelitian ini adalah akun Instagram resmi UMKM X, sebuah usaha sektor *Food and Beverage* (F&B) di kota Tembilahan, yang dipilih menggunakan teknik *purposive sampling*. Objek penelitian berupa data log historis unggahan publik rentang waktu Mei 2024 hingga Mei 2026.

Pengumpulan data dilakukan melalui dua tahap. Pertama, pilot study dilakukan terhadap 20 unggahan terbaru untuk memvalidasi indikator variabel. Kedua, pengumpulan data utama secara otomatis menggunakan *Apify Instagram Scraper* yang mengekstraksi 188 data awal publik mencakup metrik *likes*, komentar, *caption*, tanggal, serta format media. Setelah melalui tahap *preprocessing* dengan mengeliminasi data kosong (*missing values*) dan unggahan tidak relevan, diperoleh 180 data unggahan bersih (*clean data*) yang relevan untuk dianalisis.

2.2. Operasionalisasi Variabel dan Prosedur KDD

Variabel penelitian ditransformasikan ke dalam struktur matematika yang dibagi menjadi dua peran untuk menghindari bias perhitungan jarak algoritma yang dijabarkan pada tabel berikut.

Tabel 1. Operasional Variabel Penelitian Berbasis Model AIDA

No	Nama Variabel	Konsep Teoritis (AIDA)	Indikator Variabel & Skala Data	Peran Dalam K-Means	Justifikasi Literatur
1	Bahasa Promosi	<i>Desire</i>	Penggunaan gaya bahasa persuasif (emosional atau personalisasi). Skala biner: 0 = Tidak Ada, 1 = Ada.	Variabel Fitur (Feature)	Diadopsi dari aspek emosi dan personalisasi pesan model AIDA (Yadav et al., 2025).
2	<i>Call To Action</i> (CTA)	<i>Action</i>	Instruksi langsung untuk mendorong respons (misal: "Klik", "Beli"). Skala biner: 0 = Tidak Ada, 1 = Ada.	Variabel Fitur (Feature)	Diadopsi dari elemen instruksi tindakan langsung pada caption (Yadav et al., 2025).
3	Akses Transaksi	<i>Action</i>	Kehadiran fitur interaktif berupa tautan pembelian atau kontak pada caption atau bio. Skala biner: 0 = Tidak Ada, 1 = Ada.	Variabel Fitur (Feature)	Diadopsi dari kemudahan konversi tahap <i>action</i> (Rahardja et al., 2026) (Sarfraz et al., 2025).
4	Business Tag	<i>Action</i>	Penyebutan (mention/tag) akun bisnis lain atau mitra kolaborasi pada unggahan. Skala biner: 0 = Tidak Ada, 1 = Ada.	Variabel Fitur (Feature)	Diadopsi dari kemudahan konversi tahap <i>action</i> (Rahardja et al., 2026)(Nallasivam et al., 2025).
5	<i>Engagement Rate</i> (ER)	<i>Action Evaluation</i>	Rasio interaksi publik: $\frac{\text{Jumlah Likes} + \text{Komentar}}{\text{Total Pengikut}} \times 100\%$	Variabel Utama & Evaluasi	Diadopsi dari metrik evaluasi performa klaster komunikasi (Nugraha et al., 2025)(Arief et al., 2026).
6	Hari Unggah	<i>Attention</i>	Skala: Numerik (rasio) Hari publikasi konten pada Instagram (Senin s.d Minggu) Skala: Kategorikal / Nominal	Variabel Karakteristik (Kontekstual Klaster)	Diadopsi dari aspek manajemen waktu distribusi konten [5].
7	Jam Unggah	<i>Attention</i>	Waktu rilis konten berdasarkan format waktu 24 jam. Skala: Kategorikal / Interval	Variabel Karakteristik (Kontekstual Klaster)	Diadopsi dari pola waktu aktif audiens media sosial (Lubis & Iqbal, 2026)

8	Tipe Konten	<i>Attention Interest</i>	&	Format visual publikasi unggahan pada akun Instagram. Skala Kategorikal (Koding): 1 = <i>Single Photo / Image</i> 2 = <i>Carousel</i> 3 = <i>Reels / Video</i>	Variabel Karakteristik (Kontekstual Klaster)	Diadopsi berdasarkan preferensi format visual yang efektif dalam menangkap perhatian awal audiens (Nugraha et al., 2025).
---	-------------	---------------------------	---	--	--	---

Berdasarkan Tabel 1 di atas variabelnya penelitian disusun sekuensial mengacu pada hierarki model AIDA. Variabel Hari Unggah, Jam Unggah dan Tipe Konten mewaliki tahap *Attention* dan *Interest*. Tahap *Desire* dieksplorasi lewat Bahasa Promosi pada *caption*, sementara tahap Action dioperasionalkan melalui *Call to Action* (CTA), *Business Tag* dan Akses Transaksi. Sebagai evaluasi tindakan (*Action Evaluation*), metrik *Engagement Rate* (ER) digunakan untuk mengukur tingkat keterlibatan riil audiens (Arief et al., 2026).

Operasionalisasi rumus dan langkah komputasi dalam siklus KDD dijelaskan melalui persamaan (1) hingga persamaan (5) sebagai berikut.

Kalkulasi nilai *Engagement Rate* (ER) dihitung menggunakan persamaan (1) (Sarfraz et al., 2025) (Nugraha et al., 2025).

$$\text{Engagement Rate (ER)} = \frac{\text{Jumlah Likes} + \text{Jumlah Komentar}}{\text{Total Pengikut}} \times 100\% \quad (1)$$

Nilai ER menggambarkan tingkat keterlibatan audiens dalam bentuk persentase, di mana komponen pembilang terdiri dari akumulasi *likes* dan komentar pada setiap unggahan, sedangkan komponen penyebut menggunakan basis konstan total pengikut (*followers*) akun pada saat ekstraksi data log dilakukan.

Untuk menyetarakan skala antarvariabel fitur numerik, dilakukan transformasi menggunakan *Min-Max Normalization* pada Persamaan (2) (Lubis & Iqbal, 2026).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2)$$

X_{norm} adalah nilai data hasil normalisasi (rentang 0 hingga 1), X adalah nilai data asli, X_{min} adalah nilai minimum variabel dalam dataset, dan X_{max} adalah nilai maksimum variabel dalam dataset.

Pengelompokkan data menggunakan algoritma *K-Means Clustering* dihitung berdasarkan jarak *Euclidean* pada Persamaan (3) (Lubis & Iqbal, 2026).

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

$d(x, y)$ adalah jarak *Euclidean* antara objek data unggahan (x) dengan titik pusat klaster/centroid (y), x_i adalah koordinat data pada variabel ke- i , y_i adalah koordinat pada variabel ke- i , dan n adalah jumlah total variabel fitur yang diuji.

Penentuan jumlah klaster (K) optimal divalidasi menggunakan *Silhouette Coefficient* pada persamaan (4) (Nurhidayat et al., 2025).

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

$s(i)$ adalah nilai skor *Silhouette Coefficient* objek ke- i (rentang -1 hingga +1), $a(i)$ adalah rata-rata jarak objek ke- i terhadap seluruh objek dalam klasternya sendiri, dan $b(i)$ adalah jarak rata-rata objek ke- i terhadap objek di klaster tetangga terdekat.

Tahap penerapan (*deployment*) dilakukan dengan mentransformasikan centroid akhir ke dalam fungsi keputusan pada lembar kerja digital menggunakan persamaan (5).

$$\text{Status Klaster} = \arg \min_{k \in \{1, 2, \dots, K\}} (\text{Jarak}_{Ck})$$

Status klaster adalah kelompok klaster hasil penetapan untuk data unggahan baru (x), $d(x, C_k)$ adalah jarak *Euclidean* antara data baru x ke centroid klaster ke- k , dan K adalah jumlah total klaster optimal yang terbentuk.

2.3. Perangkat Pendukung Penelitian

Prosedur pengolahan dan komputasi data didukung oleh perangkat keras dan perangkat lunak dengan spesifikasi ringkas yang disajikan pada Tabel 1.

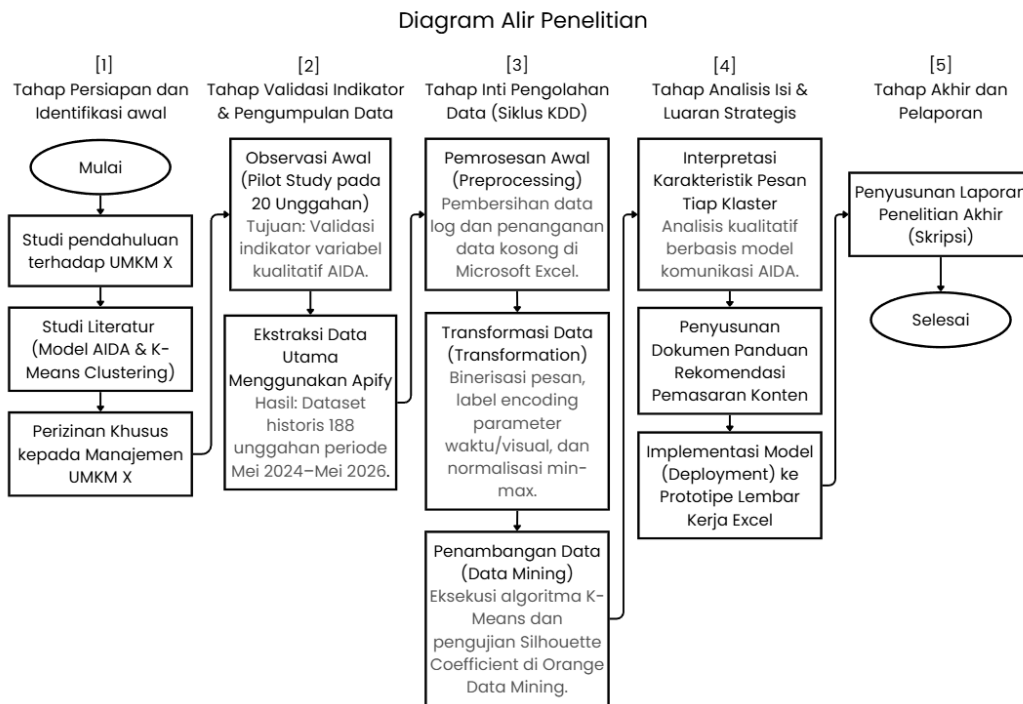
Tabel 2. Perangkat Lunak dan Perangkat Keras Pendukung

Jenis perangkat	Nama perangkat	Fungsi Operasional
Perangkat keras	Komputer/Laptop Spesifikasi Standar	Eksekusi komputasi data dan pemrosesan dokumen

Perangkat lunak	Apify (Instagram Scraper)	Ekstraksi 188 data log public historis secara otomatis
Perangkat lunak	Microsoft Excel	Preprocessing, kodifikasi biner, dan pembuatan prototipe lembar kerja
Perangkat lunak	Orange Data Mining	Ekseskusi algoritma K-Means, uji Silhouette Coefficient, ANOVA, dan Chi Square.
Perangkat lunak	Instagram & Web Browser	Verifikasi silang data riil di lapangan

2.4. Arsitektur Prosedur Penelitian

Rangkaian alur komputasi data dari tahap ekstraksi hingga menghasilkan luaran prototipe lembar kerja digital disusun ke dalam siklus KDD yang ditunjukkan pada Gambar 1.



Gambar 1. Arsitektur Prosedur Penelitian Berdasarkan Kerangka Kerja KDD

Gambar 1 menunjukkan alir yang dilalui dimulai dari tahap persiapan dan identifikasi awal, dilanjutkan dengan memvalidasi dan mengumpulkan data. Setelahnya dilanjutkan dengan pengolahan data menggunakan kerangka kerja KDD. Kemudian dilanjutkan dengan menganalisa hasil klastering dan mengimplementasikan hasil analisa ke sebuah spreadsheet sebagai sistem evaluasi mandiri.

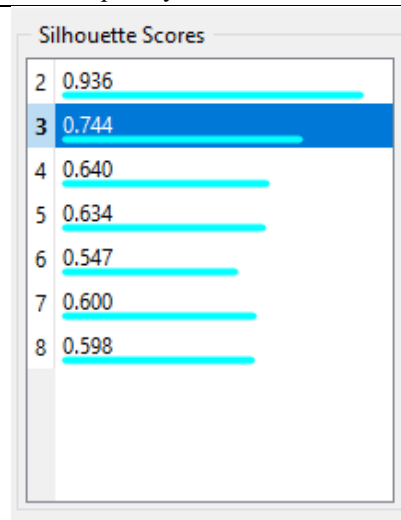
3. Hasil dan Pembahasan

Bab ini menyajikan temuan empiris hasil pengolahan data log historis akun Instagram UMKM X yang dieksekusi melalui kerangka kerja *Knowledge Discovery in Databases* (KDD). Penyajian disusun secara sistematis yang meliputi tahapan pemrosesan data, hasil pemetaan *K-Means Clustering*, pengujian validitas statistik, evaluasi pola visualisasi berbasis kerangka AIDA, hingga implementasi prototipe lembar kerja digital sebagai instrumen evaluasi mandiri.

3.1. Pemrosesan Data dan Hasil Pemetaan Kluster

Ekstraksi data historis akun Instagram UMKM X menggunakan *Apify Instagram Scraper* menghasilkan 188 data mentah. Pembersihan data (*data cleaning*) dilakukan dengan mengeliminasi data bernilai kosong (*missing values*) serta unggahan non-bisnis, sehingga menghasilkan 180 data unggahan bersih yang valid untuk dimodelkan. Ekstraksi data ini dijalankan dengan asumsi basis pengikut (*followers*) konstan saat pengambilan data log. Seluruh variabel fitur pesan yang dikodifikasi berbasis kerangka AIDA serta nilai *Engagement Rate* (ER) yang telah dinormalisasi diproses menggunakan perangkat lunak *Orange Data Mining*.

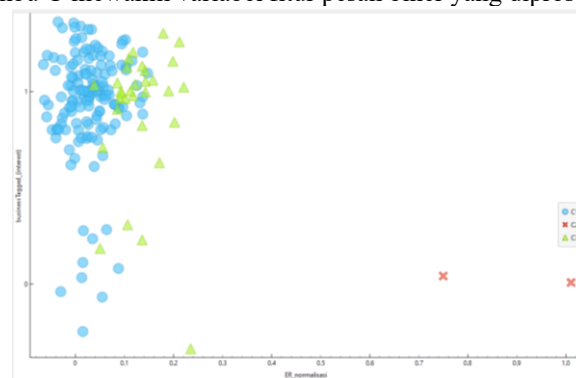
Proses pengelompokkan menggunakan algoritma *K-Means* menghasilkan tiga kluster optimal ($K=3$) sebagai struktur terbaik dengan skor *Silhouette Coefficient* sebesar 0,744. Angka ini mengindikasikan bahwa struktur kluster yang terbentuk tergolong kuat dan terpisah secara jelas. Ketiga kluster ini memetakan 180 data unggahan ke dalam variasi tingkatan performa interaksi yang terdistribusi secara sistematis.



Gambar 1. Skor Silhouette dan Box Plot

3.2. Analisis Sebaran Scatter Plot dan Box Plot

Visualisasi distribusi data hasil klasterisasi dipetakan menggunakan *Scatter Plot*, di mana sumbu X mewakili nilai ER hasil normalisasi dan sumbu Y mewakili variabel fitur pesan biner yang diproses menggunakan fitur *jittering*.

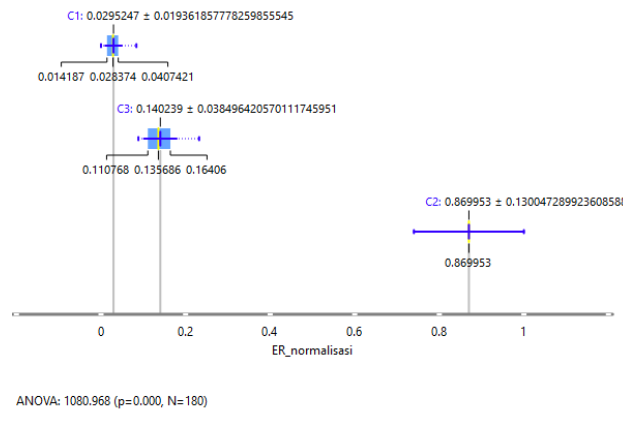


Gambar 2. Visualisasi Klaster Melalui Grafik Scatter Plot

Berdasarkan pemetaan pada variabel *businessTagged*, Klaster 1 (C1) berkumpul padat di area kiri dengan skala ER terendah (0,0–0,1). Hal ini menandakan bahwa penggunaan tag bisnis tidak memberikan dampak langsung terhadap peningkatan interaksi. Klaster 2 (C2) terisolasi di kanan bawah dengan skala ER tertinggi (0,7–1,0), sedangkan Klaster 3 (C3) tersebar di area atas dengan tingkat interaksi sedang pada skala 0,1–0,2.

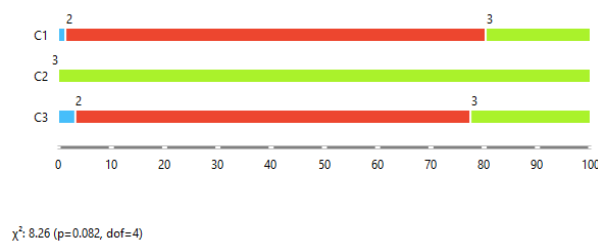
Pengujian sebaran pada variabel *CTA_Presence*, *promoLanguage*, dan *transactionAccess* menunjukkan kecenderungan pola visualisasi yang relatif homogen. Pada ketiga variabel biner tersebut, Klaster 1 terkonsentrasi di area kiri atas, yang membuktikan bahwa penyertaan kalimat ajakan (*Call to Action*), bahasa promosi (*hard-selling*), maupun petunjuk akses transaksi secara masif tidak mampu meningkatkan *Engagement Rate* audiens. Sebaliknya, Klaster 2 secara konsisten berada di pojok kanan bawah, mengindikasikan bahwa unggahan dengan performa interaksi tinggi justru tidak menonjolkan pesan promosi transaksional secara eksplisit. Sementara itu, Klaster 3 berada di area atas pada skala ER menengah (0,1–0,2), yang menunjukkan potensi interaksi sedang ketika elemen promosi disajikan secara selektif (*soft-selling*).

Untuk mempertegas perbedaan distribusi performa secara makro, analisis dilanjutkan dengan mengamati diagram *Box Plot* dan proporsi tipe konten. Klaster 1 (C1) memiliki nilai rata-rata (*mean*) ER sebesar 0,0295 dengan struktur kotak pipih di area bawah, mencerminkan homogenitas performa pada tingkat yang rendah. Klaster 3 (C3) berada pada posisi moderat dengan nilai rata-rata 0,1402 dan variabilitas sebaran yang sedikit lebih dinamis. Klaster 2 (C2) mencatatkan nilai rata-rata tertinggi sebesar 0,8699 dengan rentang garis yang memanjang, menunjukkan performa interaksi yang jauh mengungguli kelompok lainnya.



Gambar 3. Box Plot ER_Normalisasi dan Kluster

Validitas statistik atas perbedaan distribusi interaksi ini diperkuat oleh uji *Analysis of Variance* (ANOVA) pada variabel ER yang menghasilkan nilai $F = 1080,968$ ($p = 0,000$). Hal ini membuktikan secara mutlak adanya perbedaan rata-rata performa interaksi yang sangat signifikan antar-kluster.



Gambar 4. Distribusi Proporsi Tipe Konten Antar Kluster

Di sisi lain, uji Chi-Square terhadap proporsi format konten menghasilkan nilai $X^2 = 8,26$ ($p = 0,082$; $dof = 4$). Meskipun tidak signifikan secara mutlak pada taraf kepercayaan 5%, distribusi proporsi format konten secara deskriptif tetap memberikan gambaran pola penyusunan pesan yang sangat jelas.

3.3. Implementasi Prototipe Lembar Kerja Evaluasi Mandiri

Sebagai bentuk penerapan praktis dari model *centroid* yang terbentuk, dikembangkan sebuah prototipe lembar kerja digital berbasis Microsoft Excel. Alat bantu ini dirancang untuk memudahkan pemilik UMKM X dalam melakukan evaluasi dan prediksi performa konten secara mandiri tanpa memerlukan perangkat lunak komputasi yang rumit. Prosedur operasional instrumen ini diawali dengan penarikan data log historis secara berkala, yang dilanjutkan dengan penginputan metrik interaksi riil (*likes* dan *comments*) serta kodifikasi elemen pesan berbasis AIDA ke dalam sel input yang telah disediakan.

Sistem pada lembar kerja mengintegrasikan tiga tahap kalkulasi otomatis pada setiap baris data baru. Pertama, nilai *Engagement Rate* diubah ke skala $[0,1]$ menggunakan fungsi normalisasi Min-Max. Kedua, jarak geometris data baru terhadap titik pusat (*centroid*) dari ketiga kluster dihitung menggunakan formula jarak Euclidean. Ketiga, fungsi logika *Nested IF* yang dikombinasikan dengan perintah MIN digunakan untuk mengidentifikasi nilai jarak terdekat sekaligus menetapkan kategori prediktif performa konten menjadi "Rendah" (sebanding dengan Klaster 1), "Sedang" (sebanding dengan Klaster 3), atau "Tinggi" (sebanding dengan Klaster 2).

Pengujian keandalan lembar kerja dilakukan melalui dua skenario eksperimen. Skenario pertama diterapkan pada data historis aktual, di mana formula secara konsisten memberikan label klasifikasi yang padan dan identik dengan hasil pemodelan awal *Orange Data Mining*. Skenario kedua dilakukan menggunakan simulasi angka acak pada batas nilai kritis. Hasil simulasi menunjukkan bahwa struktur formula logika tetap stabil dan responsif dalam mengelompokkan data baru ke kelompok kluster terdekat berdasarkan bobot jarak Euclidean terkecil.

Pembahasan

Melalui perspektif kualitatif kerangka AIDA, temuan kuantitatif ini memperjelas efektivitas strategi komunikasi digital UMKM X. Klaster 1 mewakili pendekatan *hard-selling* yang terlalu berfokus pada tahapan *Action*. Dominasi 147 data unggahan pada klaster ini membuktikan adanya ketidakefisienan strategi pemasaran, di mana pesan transaksional yang berlebihan justru memicu kejenuhan audiens karena gagal membangun tahapan *Attention*. Sebaliknya, Klaster 2 mewakili pendekatan *Organic Interest*, di mana penggunaan format video pendek terbukti sangat efektif menstimulasi respons *Attention* dan *Interest* awal audiens melalui daya tarik visual yang

dinamis. Klaster 3 mencerminkan pendekatan *selective AIDA balance*, di mana pesan promosi disajikan secara berimbang (*soft-selling*) sehingga mampu menjaga stabilitas minat (*Desire*) audiens pada tingkat interaksi menengah.

Temuan ini mengindikasikan bahwa frekuensi unggahan yang tinggi tidak secara otomatis meningkatkan *Engagement Rate* apabila format penyampaian pesan tidak sesuai dengan preferensi audiens. Pihak manajemen UMKM X disarankan untuk mengubah orientasi pembuatan konten dari pendekatan *hard-selling* transaksional ke strategi *hybrid*, yaitu memanfaatkan video pendek sebagai penarik perhatian awal (*Attention*), yang kemudian dipadukan dengan pesan promosi selektif (*soft-selling*) untuk mendorong tindakan pembelian (*Action*). Kehadiran prototipe lembar kerja memberikan instrumen evaluasi terapan yang aplikatif bagi UMKM X untuk merencanakan dan menguji eksperimen konten mingguan secara terukur berbasis data log.

4. Kesimpulan

Implementasi algoritma *K-Means Clustering* berbasis kerangka kerja *Knowledge Discovery in Databases* (KDD) yang dipadukan dengan model hierarki komunikasi AIDA terbukti mampu mengelompokkan dan mengevaluasi 180 data unggahan historis akun Instagram UMKM X secara objektif dan valid. Pemodelan komputasional menghasilkan tiga klaster optimal dengan skor *Silhouette Coefficient* sebesar 0,744, yang didukung oleh signifikansi uji ANOVA ($F = 1080,968$; $p = 0,000$) serta gambaran deskriptif uji *Chi-Square* ($X^2 = 8,26$; $p = 0,082$). Profil performa konten UMKM X saat ini masih didominasi oleh kelompok konten *hard-selling* berformat gambar ganda (*Carousel*) pada Klaster 1 (147 data), yang secara konsisten menghasilkan rata-rata *Engagement Rate* rendah sebesar 0,0295. Profil performa konten UMKM X saat ini masih didominasi oleh kelompok konten *hard-selling* berformat gambar ganda (*Carousel*) pada Klaster 1 (147 data), yang secara konsisten menghasilkan rata-rata *Engagement Rate* rendah sebesar 0,0295. Sebaliknya, format video pendek pada Klaster 2 (2 data) mencetak rata-rata interaksi tertinggi sebesar 0,8699, disusul oleh pendekatan *soft-selling* pada Klaster 3 (31 data) yang menjaga stabilitas interaksi di tingkat menengah dengan rata-rata 0,1402. Temuan ini mempertegas bahwa kehadiran elemen pesan transaksional biner secara mandiri tidak efektif mendorong interaksi tanpa diawali stimulus visual yang tepat untuk memicu *Attention* dan *Interest* audiens.

Implikasi praktis dari penelitian ini mengarahkan UMKM X untuk mentransisikan strategi pemasarannya dari pendekatan transaksional penuh menuju strategi konten *hybrid* yang adaptif dan berimbang. Pembuat konten disarankan mengadopsi kekuatan visual dinamis format video pendek sebagai penarik perhatian utama, yang kemudian dipadukan secara selektif dengan elemen promosi guna menjaga minat (*Desire*) hingga mendorong tindakan (*Action*) audiens secara berkelanjutan. Penerapan solusi Sistem Informasi diwujudkan melalui prototipe lembar kerja digital (*spreadsheet*) yang mengintegrasikan rumus normalisasi Min-Max, jarak Euclidean, dan fungsi logika keputusan. Alat bantu ini terbukti stabil dan responsif dalam memprediksi kategori performa konten baru secara mandiri tanpa memerlukan biaya infrastruktur yang tinggi. Untuk pengembangan akademis di masa mendatang, peneliti selanjutnya disarankan untuk memperluas cakupan variabel dengan menyertakan metrik jangkauan (*reach*), impresi, serta durasi tontonan (*watch time*), sekaligus mengeksplorasi algoritma clustering lainnya dengan rentang waktu ekstraksi data yang lebih panjang guna menangkap dinamika tren perilaku audiens secara lebih presisi.

Reference

- Ambarsari, E. W., Fathudin, D., & Alfiani, G. (2024). *Utilizing K-Means Clustering to Understanding Audience Interest in SEO- Optimized Media Content*. 3(2), 208–214. <https://doi.org/10.47065/comforch.v3i2.1207>
- Arief, H., Sembiring, M. T., & Rini, E. S. (2026). *Strategy To Increase Followers and Account Reach Social Media Instagram PT. Bank SUMUT*. 9(2). <https://doi.org/10.32734/lwsa.v9i2.2845>
- Kedi, W. E., Ejimuda, C., & Idemudia, C. (2024). *Machine learning software for optimizing SME social media marketing campaigns*. 5(7), 1634–1647. <https://doi.org/10.51594/csitrj.v5i7.1349>
- Lubis, F. I., & Iqbal, M. (2026). *Analisis Waktu Optimal Unggah Instagram Menggunakan K-Means Clustering pada Data Sintetis*. 5(1), 1497–1504. <https://doi.org/https://doi.org/10.31004/riggs.v5i1.6042>
- Nafsyah, A. S., Maulidiah, S. R., Nurlia, A. S., & Putri, W. (2022). *ANALISIS KONTEN MEDIA SOSIAL INSTAGRAM BY . U SEBAGAI MEDIA PENYEBARAN INFORMASI & KOMUNIKASI Digital Public Relations , Universitas Telkom Digital Public Relations , Universitas Telkom*. 2(2), 1–11. <https://doi.org/https://doi.org/10.35842/massive.v2i2.57>
- Nallasivam, A., Desai, G., Bandhu, D., & Jagadeeswari, I. U. (2025). *Predicting Consumer Adoption of Luxury Products via Instagram Marketing : A Machine Learning Approach Anitha Nallasivam Guruprasad Desai Indonesia*. 1–39. <https://f1000research.com/articles/14-890>
- Nasution, A. (2021). *STRATEGI PEMASARAN BISNIS KULINER MENGGUNAKAN INFLUENCER MELALUI MEDIA SOSIAL INSTAGRAM. Jurnal Bisnis Corporate*, 6(1), 11–18. <https://doi.org/https://doi.org/10.46576/jbc.v6i1.1484>
- Nugraha, D. A., Priandono, T. E., Ainiyyah, N. F., & Ilham, Y. (2025). *Analisis Isi Konten Media Sosial Instagram Pemerintah Indonesia di Tingkat Kementerian dan Lembaga Content Analysis of the Indonesian Government ' s Instagram Accounts at the Ministry and Agency Levels*. 9(1), 42–52. <https://doi.org/https://doi.org/10.56873/jimik.v9i1.532>
- Nurhidayat, ., Fadillah, N., Syahrul, Dillah, S., & Surianto, D. F. (2025). *K-Means-Based Behavioral Segmentation of Social Media Users for Digital Communication Analysis in Indonesia*. 5(2), 249–261. <https://doi.org/10.24042/ijecs.v5i2.28749>
- Rahardja, U., Aini, Q., Supriati, R., & Al Hafiz, M. A. (2026). *Optimizing Digital Promotional Graphic Design Strategies Using the AIDA Model*. 7(2), 220–230. <https://doi.org/https://doi.org/10.34306/ajri.v7i2.1425>
- Rahman, A., Kusumadewi, R. A., & Nurzaman, F. (2025). *STRATEGI PEMASARAN DIGITAL DALAM MENINGKATKAN DAYA*

- SAING PERUSAHAAN DI ERA TRANSFORMASI TEKNOLOGI. *Idarah: Jurnal Manajemen Pemasaran*, 2(1), 8–20.
<https://doi.org/https://doi.org/10.70283/idarrah.v2i1.64>
- Salsabila, W. (2025). *PENGARUH SOFT SELLING CONTENT TERHADAP PURCHASE INTENTION YANG DIMEDIASI EMOTIONAL ENGAGEMENT (Studi pada Brand Ecinis di Pulau Jawa)*. <http://etheses.uin-malang.ac.id/id/eprint/83214>
- Sarfraz, M., Kurdi, B. Al, & Rafiq, M. (2025). *How Digital Marketing Shapes Consumer Decision-Making Employing (AIDA) Model with Respect to Consumer Knowledge and Consumer Experience*. 39–48.
- Yadav, R., Tripathi, S., & Shukla, S. S. (2025). *AIDA Model For Social Media Marketing*. 01(02), 1–4.
<https://jdacm.com/index.php/jdacm/article/view/9>
- Yanti, A. R. (2023). STRATEGI KOMUNIKASI PEMASARAN BRAND SKINCARE MAKO BY SERIS DALAM MENINGKATKAN ENGAGEMENT MELALUI MEDIA SOSIAL. *Journal of Management, Education, and Communication*, 1(3).
<https://doi.org/https://doi.org/10.64685/jmec.v1i2.11>