

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Pemasaran Digital dan Peran TikTok sebagai Kanal Promosi

Pemasaran digital adalah kegiatan pemasaran yang memanfaatkan media digital dan internet untuk menjangkau, berinteraksi, serta membangun hubungan dengan konsumen secara terukur dan *real-time* [20], [21]. Media sosial menjadi kanal utama karena memungkinkan distribusi konten, interaksi dua arah, dan pemanfaatan data analitik untuk optimasi kampanye [20], [22].

TikTok sebagai platform video pendek menawarkan potensi besar melalui algoritma *For You Page* (FYP), fitur *live*, kolaborasi dengan *influencer*, serta iklan berbayar sehingga efektif meningkatkan *brand awareness*, interaksi audiens, dan konversi penjualan, termasuk bagi UMKM [20], [21], [22], [23]. Keberhasilan strategi promosi di TikTok sangat dipengaruhi kreativitas konten, relevansi budaya, dan pemanfaatan insight data performa konten [20], [22], [23].

2.1.2 Metrik Performa Konten Video TikTok

Performa konten video TikTok dapat diukur melalui serangkaian metrik kuantitatif yang mencerminkan jangkauan (*reach*), keterlibatan (*engagement*), dan dampak audiens. Metrik utama meliputi: (1) *Video Views*: jumlah kali video ditonton, sebagai indikator jangkauan organik; (2) *Likes*: bentuk apresiasi sederhana yang mencerminkan penerimaan konten; (3) *Comments*: indikator keterlibatan mendalam dan potensi diskusi; (4) *Shares*: metrik viralitas yang

menunjukkan kesediaan audiens menyebarkan konten; serta (5) *Profile Views*: indikator ketertarikan audiens untuk mengenal brand lebih lanjut [9], [16], [24].

Untuk menormalisasi interaksi terhadap jangkauan, *Engagement Rate* dihitung dengan rumus:

$$ER = \frac{L + C + S}{V} \times 100\%$$

Keterangan:

$L = Likes$

$C = Comments$

$S = Shares$

$V = Video Views$

Metrik komposit ini memungkinkan perbandingan performa konten yang lebih adil, terlepas dari perbedaan skala *views* [16], [25]. Penelitian oleh Simanjuntak et al. menemukan bahwa interaktivitas pesan dan waktu posting berpengaruh signifikan terhadap *customer engagement* di TikTok, sedangkan Wiguna et al. menekankan pentingnya mempertimbangkan bobot relatif antar metrik dalam evaluasi konten [24]. Dalam penelitian ini, kelima metrik tersebut akan digunakan sebagai variabel input dalam proses klusterisasi untuk mengidentifikasi pola performa konten secara holistik

2.1.3 Segmentasi Data dalam Pemasaran Digital

Segmentasi merupakan proses membagi pasar atau data konsumen ke dalam kelompok-kelompok homogen berdasarkan karakteristik tertentu seperti demografi, perilaku, atau nilai ekonomi, agar strategi pemasaran dapat lebih tepat sasaran [24], [25], [26]. Di era digital, segmentasi tidak hanya dilakukan pada

pelanggan, tetapi juga pada audiens maupun konten, misalnya segmentasi pelanggan *e-commerce* atau pengguna media sosial [24], [25].

Dalam pemasaran digital, segmentasi berbasis data memungkinkan identifikasi kelompok “pelanggan terbaik”, pelanggan baru, maupun pelanggan tidak aktif sehingga tiap segmen dapat diperlakukan dengan strategi komunikasi berbeda [25], [26]. Teknik segmentasi berbasis data mining dan pembelajaran mesin, termasuk *clustering*, banyak digunakan untuk mengungkap pola perilaku dan preferensi tersembunyi yang sulit diidentifikasi secara manual [20], [24], [26], [27].

2.1.4 Data Mining dan Pembelajaran Tanpa Supervisi (*Unsupervised Learning*)

Data mining adalah proses mengekstraksi pola, hubungan, dan pengetahuan yang berguna dari kumpulan data besar dengan bantuan metode statistik, *machine learning*, dan basis data [24], [25], [26]. Salah satu kategori utama dalam data mining adalah *unsupervised learning*, yaitu pembelajaran tanpa label kelas, yang bertujuan menemukan struktur atau kelompok alami dalam data [24], [28].

Dalam konteks pemasaran digital, *unsupervised learning* banyak dimanfaatkan untuk segmentasi pelanggan, pengelompokan perilaku pengguna, analisis pola penggunaan media sosial, dan pengelompokan konten [20], [24], [27]. Berbeda dengan *supervised learning* yang membutuhkan label (misalnya “beli” atau “tidak beli”), *unsupervised learning* bekerja hanya berdasarkan kesamaan fitur sehingga cocok ketika label tidak tersedia atau mahal diperoleh [24], [28].

2.1.5 Algoritma *K-Means Clustering*

K-Means Clustering adalah algoritma *partitional clustering* yang mengelompokkan nn objek ke dalam kk kluster dengan meminimalkan *within-cluster sum of squares* (WCSS) [12], [15]. Prinsip kerjanya meliputi empat langkah iteratif: (1) inisialisasi kk centroid secara acak atau menggunakan *k-means++* untuk stabilitas lebih baik; (2) penugasan setiap objek ke *centroid* terdekat berdasarkan jarak *Euclidean*; (3) pembaruan posisi *centroid* sebagai rata-rata objek dalam kluster; dan (4) pengulangan langkah 2–3 hingga konvergensi (perubahan *centroid* $< threshold$) [21], [23].

Rumus jarak *Euclidean* antara objek x_i dan centroid c_j didefinisikan sebagai:

$$d(x_i, c_j) = \sqrt{\sum_{p=1}^m (x_{ip} - c_{jp})^2}$$

Keterangan:

x_i = objek ke- i

c_j = centroid ke- j

m = jumlah fitur

Keunggulan K-Means meliputi efisiensi komputasi, kemudahan interpretasi, dan skalabilitas untuk *dataset* menengah-besar [15]. Namun, algoritma ini sensitif terhadap inisialisasi *centroid*, *outlier*, dan skala variabel, sehingga memerlukan *preprocessing* yang cermat [12], [14]. Harani et al. berhasil

mengimplementasikan K-Means berbasis Python untuk segmentasi pelanggan *IndiHome*, membuktikan kelayakan algoritma ini dalam konteks bisnis digital Indonesia [17].

2.1.6 Penentuan Jumlah Kluster Optimal

Karena jumlah kluster k harus ditentukan sebelum eksekusi algoritma, peneliti umumnya menggunakan pendekatan *heuristik* seperti *Elbow Method* dan *Silhouette Analysis* [14], [15]. *Elbow Method* memplot nilai WCSS terhadap variasi k ; titik “siku” (*elbow*) menunjukkan penambahan k yang tidak lagi memberikan penurunan WCSS signifikan, sehingga dianggap sebagai k optimal [12]. Namun, metode ini bersifat subjektif dan dapat ambigu pada *dataset* kompleks.

Silhouette Analysis memberikan ukuran objektif dengan menghitung koefisien siluet $s(i)$ untuk setiap objek:

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

Keterangan:

$a(i)$ = rata-rata jarak intra-kluster

$b(i)$ = rata-rata jarak ke kluster terdekat lainnya

Nilai siluet berkisar antara -1 hingga 1; nilai >0.50 menunjukkan struktur kluster yang jelas, sedangkan <0.25 mengindikasikan tumpang tindih antar kluster [14]. Dalam penelitian ini, kombinasi kedua metode akan digunakan untuk menentukan k optimal, dengan pertimbangan interpretasi bisnis terhadap jumlah segmen yang *actionable*.

2.1.7 Evaluasi Kualitas Klaster

Setelah klaster terbentuk, kualitas segmentasi perlu divalidasi menggunakan metrik internal yang tidak memerlukan label eksternal [12], [15]. Tiga indeks yang umum digunakan adalah:

1. *Silhouette Score*: Rata-rata koefisien siluet seluruh objek; semakin mendekati 1, semakin baik pemisahan dan kepadatan klaster.
2. *Davies-Bouldin Index* (DBI): Mengukur rasio dispersi intra-klaster terhadap jarak antar-klaster; nilai lebih kecil menunjukkan klaster yang lebih terpisah jelas.
3. *Calinski–Harabasz Index* (CHI): Rasio antara *between-cluster dispersion* dan *within-cluster dispersion*; nilai lebih tinggi mengindikasikan segmentasi yang lebih baik [14], [15].

Aditya et al. menggunakan ketiga metrik tersebut untuk memvalidasi klaster preferensi konten TikTok, memastikan bahwa segmentasi yang dihasilkan tidak hanya statistik valid tetapi juga bermakna secara bisnis. Dalam penelitian ini, ambang batas penerimaan ditetapkan: *Silhouette Score* ≥ 0.50 , DBI < 1.0 , dan CHI > 100 , sebagai acuan objektif kualitas klaster sebelum interpretasi profil dan perumusan rekomendasi [29].

2.1.8 Tahapan *Preprocessing* Data untuk Clustering

Kualitas hasil clustering sangat bergantung pada kualitas data input. Oleh karena itu, tahapan *preprocessing* menjadi krusial dan meliputi [12], [15]:

1. *Data Cleaning*: Menangani *missing values* (imputasi median/mode atau penghapusan jika $<5\%$), mengoreksi nilai anomali (misal: *comments* negatif), dan menghapus duplikasi observasi.
2. *Feature Selection*: Memilih variabel numerik yang relevan dengan tujuan penelitian; dalam kasus ini: *Video_Views*, *Likes*, *Comments*, *Shares*, dan *Engagement_Rate*.
3. *Feature Engineering*: Membuat variabel komposit seperti *Engagement_Rate* dan *New_Viewers_Ratio* untuk memperkaya representasi performa konten.
4. *Outlier Treatment*: Mendeteksi dan menangani pencilan menggunakan metode IQR (*Interquartile Range*) atau Z-score, karena K-Means sensitif terhadap nilai ekstrem [23].
5. *Normalization/Scaling*: Menerapkan *StandardScaler* atau *MinMaxScaler* untuk menstandarisasi skala variabel, mengingat K-Means menggunakan jarak *Euclidean* yang bias terhadap fitur berskala besar [12], [15].

Harani et al. menekankan bahwa tanpa normalisasi, variabel seperti *Video_Views* (ribuan) akan mendominasi perhitungan jarak, sehingga kluster hanya mencerminkan perbedaan jangkauan, bukan pola interaksi yang holistik [17]. Oleh karena itu, seluruh tahapan *preprocessing* akan didokumentasikan secara rinci di Bab III untuk memastikan *reproducibility* penelitian.

2.1.9 Optimasi Strategi Pemasaran Digital Berbasis Data

Hasil segmentasi kluster tidak berhenti pada deskripsi statistik, melainkan harus diterjemahkan menjadi rekomendasi strategis yang *actionable* [4], [8]. Setiap profil kluster dianalisis berdasarkan rata-rata dan distribusi fitur, kemudian

dipetakan ke dimensi strategi pemasaran: (1) Format Konten: durasi, gaya visual, elemen interaktif; (2) Waktu Unggah: jam dan hari dengan engagement tertinggi per klaster; (3) Pesan & CTA: penyesuaian nada, nilai informasi, dan ajakan bertindak; serta (4) Alokasi Promosi: prioritas budget untuk klaster dengan potensi ROI tertinggi [9], [17], [18].

Pendekatan ini selaras dengan prinsip *data-driven decision making*, di mana keputusan strategis didasarkan pada bukti empiris, bukan intuisi [10], [11]. Angelin & Paramita menunjukkan bahwa brand kecantikan yang menyesuaikan konten dengan preferensi segmen audiens di TikTok mengalami peningkatan *engagement rate* hingga $3,2\times$ [1]. Dalam konteks Perusahaan Kecap Djoe-Hoa, rekomendasi akan difokuskan pada optimasi konten promosi produk kecap, dengan mempertimbangkan karakteristik audiens lokal, tren kuliner, dan nilai budaya yang relevan [7], [8]. Validasi rekomendasi dapat dilakukan melalui *expert review* dengan tim pemasaran atau diskusi terbatas sebelum implementasi skala penuh.

2.1.10 Pengambilan Keputusan Berbasis Data (*Data-Driven Decision Making*)

Pengambilan keputusan berbasis data (*Data-Driven Decision Making/DDDM*) adalah pendekatan manajerial di mana keputusan strategis tidak lagi didasarkan pada intuisi atau pengalaman subjektif semata, melainkan pada analisis data empiris dan pola yang terukur [30], [31]. Dalam konteks pemasaran digital, penerapan DDDM memungkinkan perusahaan untuk mengidentifikasi pola perilaku audiens secara objektif, meminimalkan risiko kegagalan kampanye, dan mengoptimalkan alokasi sumber daya. Penelitian ini mengadopsi prinsip DDDM

dengan menggunakan algoritma K-Means untuk menerjemahkan data mentah metrik TikTok menjadi segmen performa yang dapat ditindaklanjuti (*actionable*), sehingga pergeseran dari metode trial-and-error menuju strategi berbasis bukti (*evidence-based*) dapat terwujud [32].

2.2 Penelitian Terdahulu

Penelitian terdahulu menjadi landasan penting untuk mengidentifikasi celah pengetahuan (*research gap*) dan memposisikan kontribusi orisinalitas penelitian ini. Berbagai kajian terkini telah mengeksplorasi peran TikTok dalam pemasaran digital, penerapan *data mining*, serta penggunaan algoritma *clustering* untuk segmentasi audiens atau analisis konten. Namun, sebagian besar studi masih berfokus pada tinjauan konseptual, segmentasi pelanggan, atau analisis kualitatif, tanpa secara eksplisit mengintegrasikan metrik performa konten video ke dalam model segmentasi berbasis *unsupervised learning* untuk keperluan optimasi strategi pemasaran. Untuk memperjelas posisi dan kebaruan (*novelty*) penelitian, berikut disajikan matriks tinjauan literatur yang membandingkan sepuluh studi relevan dalam kurun waktu 2023–2025, mencakup aspek metode, hasil, serta perbedaan mendasar dengan penelitian yang dilakukan.

Penerapan Metode *K-Means Clustering* Untuk Segmentasi Performa Konten Video Tiktok Sebagai Dasar Optimasi Strategi Pemasaran Digital

Tabel 2. 1 Matriks literatur review dan posisi penelitian

No	Judul Penelitian	Nama Peneliti, Tahun, Index	Metode Penelitian	Hasil	Keunggulan dan Kelemahan	Perbandingan
1	<i>The Impact of Short Video Content and</i>	Islam & Kabir, 2025,	<i>Systematic Literature Review</i>	<i>Short video + personalisasi</i>	Keunggulan: Pemetaan komprehensif tren short	Penelitian ini menerapkan K-Means secara empiris

No	Judul Penelitian	Nama Peneliti, Tahun, Index	Metode Penelitian	Hasil	Keunggulan dan Kelemahan	Perbandingan
	<i>Social Media Influencers on Digital Marketing Success: A Systematic Literature Review of Smartphone Usage</i>	<i>Innovatech Engineering Journal</i> (SJR/Scopus) [22]	(78 artikel, 2018–2024)	algoritmik meningkatkan <i>engagement</i> ~60%; metrik kunci: <i>retention, share, real-time analytics</i>	video marketing & metrik performa. Kelemahan: Tidak ada implementasi model <i>clustering</i> ; tidak spesifik K-Means	pada data performa konten TikTok riil untuk membentuk segmen performa, lalu mengusulkan strategi optimasi berbasis data, mengisi gap implementasi praktis di level konten
2	<i>Unpacking Multimodal Fact-Checking: Features and Engagement of Fact-Checking Videos on Chinese TikTok (Douyin)</i>	Lu & Shen, 2023, <i>Social Media + Society</i> (Scopus Q1) [33]	<i>Hybrid automated & manual content analysis; feature clustering (unsupervised)</i> pada 4.309 video	Mengidentifikasi 3 tipe video <i>fact-checking</i> ; fitur audiovisual & strategi persuasif terkait <i>likes, komentar, reshare</i>	Keunggulan: <i>Dataset</i> besar & fokus <i>engagement</i> ; menggunakan <i>clustering fitur</i> . Kelemahan: Bukan konteks pemasaran/brand; tidak menjelaskan segmentasi performa untuk strategi	Mengadopsi semangat pengelompokan berbasis <i>fitur+engagement</i> , tetapi objek penelitian ini adalah konten pemasaran brand dan outputnya berupa segmen performa untuk rekomendasi strategi pemasaran digital
3	<i>Typology and Impact of Follower Responses to</i>	Rynarzewska et al., 2024,	<i>Mixed methods, netnografi + K-</i>	Tiga klaster respons <i>follower</i>	Keunggulan: Menggunakan K-Means	Tetap memakai K-Means, tetapi unit analisis

No	Judul Penelitian	Nama Peneliti, Tahun, Index	Metode Penelitian	Hasil	Keunggulan dan Kelemahan	Perbandingan
	<i>Social Media Influencer “Promotional Crisis”</i>	<i>Psychology & Marketing</i> (Scopus Q1) [34]	Means <i>cluster analysis</i> pada komentar TikTok	memengaruhi persepsi kredibilitas & brand saat krisis	pada data TikTok; menjelaskan perilaku komunitas. Kelemahan: Fokus krisis & komentar, bukan performa konten rutin & strategi konten	penelitian ini adalah performa konten video (metrik objektif), bukan respons <i>follower</i> saat krisis, dengan tujuan optimasi kampanye rutin
4	<i>Novel Application of K-Means Algorithm for Unique Sentiment Clustering in 2024 Korean Movie Reviews on TikTok Platform</i>	Erdiani et al., 2025, MATRIK (Sinta/Scopus) [35]	<i>Web crawling</i> komentar, IndoBERT, <i>SentimentIntensity Analyzer</i> , K-Means + PCA	Distribusi sentimen per genre film; action lebih terpolarisasi, <i>romance/comedy</i> cenderung netral	Keunggulan: Implementasi penuh K-Means di TikTok; teks Indonesia relevan. Kelemahan: Objek: review film, bukan konten pemasaran; keluaran tidak dikaitkan strategi marketing	Mengembangkan penggunaan K-Means dari <i>sentiment clustering</i> ke segmentasi performa konten pemasaran berbasis metrik kuantitatif, dan secara eksplisit menghubungkannya ke rekomendasi strategi
5	<i>Personalized Marketing Strategy in Digital Business Using Data Mining Approach</i>	Yusnidar et al., 2023, IJSECS (Scopus) [36]	<i>Mixed methods; data mining multi-platform</i> (WhatsApp, IG,	Menemukan segmen pelanggan & pola rekomendasi produk; menekankan	Keunggulan: Memadukan data mining & strategi pemasaran; menyentuh TikTok	Berfokus pada profil performa konten TikTok (bukan pelanggan) dengan K-Means; memperkaya

No	Judul Penelitian	Nama Peneliti, Tahun, Index	Metode Penelitian	Hasil	Keunggulan dan Kelemahan	Perbandingan
			TikTok, Shopee); segmentasi & pola pembelian	keseimbangan personalisasi–privasi	sebagai kanal. Kelemahan: Segmentasi di level pelanggan, bukan konten; tidak spesifik K-Means atau metrik video	pendekatan data mining dengan unit analisis dan variabel metrik video yang berbeda
6	<i>Sustainable Marketing and the Role of Social Media: An Experimental Study Using NLP</i>	Dash et al., 2023, <i>Sustainability</i> (Scopus Q1) [37]	<i>Text mining</i> + K-Means clustering pada korpus artikel (NLP, TF-IDF, LSA)	Menghasilkan klaster tema riset utama (<i>green marketing, sustainable SMM, influencer marketing</i>)	Keunggulan: Menunjukkan penggunaan K-Means untuk identifikasi tren; menghubungkan <i>social media</i> & marketing. Kelemahan: Data artikel ilmiah, bukan data performa <i>platform</i> ; tidak menyoroti rekomendasi operasional	Memindahkan K-Means dari <i>clustering</i> tema teks ilmiah ke clustering performa konten video TikTok; memberikan kontribusi praktis langsung bagi brand, bukan hanya peta topik akademis
7	<i>A Review on Customer Segmentation Methods for Personalized</i>	Gomes & Meisen, 2023, <i>Information</i>	Systematic review (105 publikasi), analisis metode	Menyusun kerangka 4 fase (<i>data collection,</i>	Keunggulan: Landasan teori kuat untuk K-Means; menjelaskan <i>pipeline</i>	Mengadopsi kerangka 4 fase tersebut, tetapi mengganti objek dari

No	Judul Penelitian	Nama Peneliti, Tahun, Index	Metode Penelitian	Hasil	Keunggulan dan Kelemahan	Perbandingan
	<i>Customer Targeting in E-Commerce Use Cases</i>	<i>Systems and e-Business Management</i> (Scopus Q1) [38]	segmentasi; menyoroti dominasi K-Means	<i>representation, segmentation, targeting</i>); K-Means paling sering dipakai	lengkap dari data ke targeting. Kelemahan: Fokus <i>e-commerce & customer segmentation</i> , bukan konten video TikTok	pelanggan <i>e-commerce</i> ke konten TikTok; mengisi celah penerapan K-Means pada segmentasi performa konten sebagai dasar targeting strategi
8	<i>Decoding Consumer Behaviour: Leveraging Big Data and Machine Learning for Personalized Digital Marketing</i>	Mohammad et al., 2025, <i>Data and Metadata</i> (Scopus) [32]	<i>Mixed-methods; big data analytics</i> , K-Means pada data perilaku & emosi, <i>sentiment analysis</i> , regresi	Tiga segmen konsumen berdasar <i>engagement & sentimen</i> ; sentimen tidak signifikan terhadap konversi	Keunggulan: Mengombinasikan segmen perilaku & emosi; menggunakan K-Means dalam digital marketing. Kelemahan: Tidak spesifik TikTok/video; fokus <i>purchase conversion</i> , bukan performa konten per video	Mengisi gap dengan fokus khusus pada performa konten TikTok dan menjadikan hasil kluster sebagai dasar rekomendasi strategi konten, bukan langsung ke konversi; memberi insight taktis operasional
9	<i>A Digital Marketing Strategy Analysis and Application in the New</i>	Liu, 2025, <i>Advances in Economics</i> ,	Studi kasus, analisis strategis (<i>Porter's Five</i>	Menjelaskan keberhasilan TikTok melalui algoritma	Keunggulan: Fokus TikTok & strategi digital marketing; menggunakan	Melengkapi level strategis <i>platform</i> dengan analisis operasional di level konten

No	Judul Penelitian	Nama Peneliti, Tahun, Index	Metode Penelitian	Hasil	Keunggulan dan Kelemahan	Perbandingan
	<i>Media Era: A Case Study of TikTok</i>	<i>Management and Political Sciences</i> [39]	<i>Forces, AARRR, Yip framework, user behavior data)</i>	rekomendasi & user stickiness; menawarkan strategi optimasi berbasis segmentasi pengguna	data perilaku pengguna. Kelemahan: Tidak menyajikan model <i>clustering</i> eksplisit; tidak mengelompokkan performa konten sebagai basis rekomendasi	(klaster performa akun tertentu) untuk memberi rekomendasi konkrit kepada pelaku usaha/brand
10	<i>Projected Image of Taquile Island–Peru: An Unsupervised Learning Approach with AI-Supported Video Analytics of Promotional Content on TikTok</i>	Turpo et al., 2025, <i>Journal of Travel & Tourism Marketing</i> (Scopus Q1) [40]	<i>Topic modeling + image clustering (unsupervised)</i> pada 51 video promosi wisata di TikTok (1.301 frame)	Mengungkap tema dominan & pola visual (keindahan alam, tradisi, kehidupan komunitas) untuk kampanye destinasi autentik	Keunggulan: Langsung menganalisis konten promosi di TikTok; menggunakan <i>unsupervised learning</i> pada visual. Kelemahan: Tidak memakai K-Means secara eksplisit; tidak memetakan segmen berdasarkan performa (<i>views/engagement</i>)	Menggabungkan pendekatan <i>unsupervised</i> pada konten TikTok dengan K-Means dan fokus pada performa kuantitatif, memberikan ukuran segmentasi yang terukur dan implikasi strategisnya bagi brand

Sumber: Diolah dari penelitian terdahulu (2023–2025)

Berdasarkan matriks tinjauan literatur di atas, dapat disimpulkan bahwa meskipun berbagai penelitian telah mengeksplorasi pemasaran TikTok, penerapan *unsupervised learning*, dan analisis konten media sosial, belum ada studi yang secara spesifik mengklasterisasi performa konten video TikTok berbasis metrik objektif (*views, likes, comments, shares, engagement rate*) untuk keperluan optimasi strategi pemasaran digital brand lokal. Penelitian ini menempati posisi unik dengan tiga kontribusi utama: (1) Objek baru, yaitu segmentasi performa konten video TikTok (bukan pelanggan, sentimen komentar, atau tema akademik); (2) Metode terstruktur, penerapan *K-Means Clustering* yang dikombinasikan dengan *Elbow Method* dan *Silhouette Analysis* langsung pada *dataset* metrik performa riil dari TikTok Studio; serta (3) Pendekatan aplikatif, yang menerjemahkan profil klaster menjadi rekomendasi strategi pemasaran digital yang operasional, terukur, dan kontekstual bagi pelaku usaha. Dengan demikian, penelitian ini mengisi kesenjangan antara analisis data *machine learning* dengan kebutuhan praktis pengambilan keputusan konten, memberikan kerangka *data-driven* yang dapat diadopsi oleh UMKM dan brand dalam meningkatkan efektivitas kampanye digital di TikTok.