

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi telah menggeser paradigma pemasaran digital dari komunikasi satu arah menjadi interaksi berbasis data yang terukur dan *real-time*. Di Indonesia, media sosial menjadi kanal strategis utama, dengan TikTok mencatat lebih dari 160 juta pengguna aktif pada kuartal I 2025 (berdasarkan data DataReportal/Kominfo, 2025), menjadikannya platform dengan tingkat *engagement* tertinggi dibandingkan media sosial lainnya [1], [2], [3]. Algoritma rekomendasi *For You Page* (FYP) dan fitur interaktif seperti komentar, berbagi (*share*), dan *live* memungkinkan brand membangun kedekatan dengan audiens secara dinamis [4], [5], [6]. Berbagai studi empiris juga menegaskan bahwa TikTok efektif meningkatkan *brand awareness*, interaksi audiens, dan minat beli, terutama ketika konten dikemas secara kreatif dan relevan dengan budaya lokal [7], [8], [9].

Meskipun potensi strategis tersebut telah terbukti, efektivitas kampanye sangat bergantung pada kemampuan menganalisis pola performa konten secara objektif. Perusahaan Kecap Djoe-Hoa sebagai salah satu pelaku usaha yang telah memanfaatkan TikTok untuk promosi produk, saat ini masih menghadapi tantangan dalam optimalisasi strategi kontennya. Berdasarkan observasi awal, perusahaan telah rutin mengunggah video pemasaran, namun belum memiliki pendekatan analitik berbasis data untuk mengevaluasi pola performa harian maupun akumulasi interaksi konten. Keputusan mengenai format video, waktu unggah, dan alokasi fokus kreatif masih bersifat *trial-and-error* atau mengandalkan preferensi subjektif

tim konten. Akibatnya, terdapat disparitas performa yang tidak terpetakan: konten dengan *views* tinggi namun interaksi rendah, atau sebaliknya, sehingga potensi *Return on Content Investment* (ROCI) belum maksimal dan strategi pemasaran sulit diukur keberhasilannya secara sistematis.

Menghadapi algoritma yang terus berubah dan volume data interaksi (*views*, *likes*, *comments*, *shares*) yang meningkat, pengambilan keputusan berbasis intuisi menjadi kurang memadai [10], [11], [12]. Diperlukan metode segmentasi yang mampu mengelompokkan pola performa secara otomatis untuk mendukung perumusan strategi yang tepat sasaran. *K-Means Clustering* merupakan algoritma *unsupervised learning* yang terbukti efektif mempartisi data numerik ke dalam kluster homogen berdasarkan kemiripan karakteristik [13], [14]. Penerapannya pada data analitik platform memungkinkan identifikasi segmen performa konten (misalnya: kluster *high engagement*, *viral potential*, atau *low performance*) tanpa memerlukan label awal, sehingga sangat cocok untuk mengeksplorasi pola tersembunyi dalam data historis akun secara efisien [8], [15].

Tinjauan literatur menunjukkan bahwa meskipun berbagai penelitian telah menguji pengaruh elemen konten TikTok terhadap *engagement* menggunakan metode statistik klasik atau survei [6], [16], [17], [18], serta studi yang menerapkan K-Means untuk segmentasi audiens atau persepsi [8], [15], penelitian yang secara eksplisit mengintegrasikan metrik performa konten harian ke dalam model segmentasi *unsupervised learning* sebagai dasar rekomendasi strategis masih relatif terbatas. Padahal, pemetaan kluster performa berpotensi menjadi fondasi objektif untuk menyusun panduan operasional pemasaran digital, seperti penyesuaian durasi

video, optimasi jadwal unggah, penentuan *call-to-action*, dan alokasi fokus produksi konten [4], [7], [19].

Berdasarkan uraian tersebut, terdapat kebutuhan mendesak untuk menerapkan pendekatan *data mining* guna mengatasi ketidakterukuran strategi konten di Perusahaan Kecap Djoe-Hoa. Oleh karena itu, penelitian ini mengangkat judul **“Penerapan Metode *K-Means Clustering* untuk Segmentasi Performa Konten Video TikTok sebagai Dasar Optimasi Strategi Pemasaran Digital”**.

Penelitian ini bertujuan untuk: (1) mempersiapkan dan memproses data analitik harian menjadi fitur yang siap diklusterisasi, (2) menentukan jumlah kluster optimal dan memetakan profil karakteristik setiap segmen performa, serta (3) menerjemahkan hasil segmentasi menjadi rekomendasi strategi pemasaran digital yang terukur, kontekstual, dan dapat diimplementasikan langsung oleh perusahaan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana tahapan *preprocessing* data agregat harian performa akun TikTok (pembersihan, rekayasa fitur, dan normalisasi) agar siap diolah menggunakan metode *K-Means Clustering*?
2. Berapa jumlah kluster (k) optimal yang dihasilkan dari penerapan *K-Means Clustering* pada data performa harian akun TikTok?
3. Apa saja karakteristik dan profil dari setiap kluster performa harian yang terbentuk berdasarkan metrik jangkauan dan interaksi konten?

4. Bagaimana hasil segmentasi klaster dapat divalidasi dan diterjemahkan menjadi rekomendasi optimasi strategi pemasaran digital bagi Perusahaan Kecap Djoe-Hoa?

1.3 Batasan Masalah

Agar penelitian terfokus dan konsisten, penelitian ini dibatasi pada hal-hal berikut:

1. Data yang digunakan adalah data sekunder agregat harian yang diekspor dari TikTok Studio akun pemasaran Perusahaan Kecap Djoe-Hoa (April 2025–April 2026). Unit analisis adalah performa harian akun (akumulasi metrik konten per hari), bukan data tingkat *per-video* atau *per-user session*.
2. Fitur yang digunakan meliputi *Video_Views*, *Likes*, *Comments*, *Shares*, serta *Engagement_Rate* (dihitung: $(Likes + Comments + Shares) / Video_Views \times 100$). Variabel *Profile_Views* dan rasio *New_Viewers/Total_Viewers* bersifat opsional dan hanya digunakan jika uji korelasi awal menunjukkan relevansi.
3. Segmentasi hanya menggunakan algoritma *K-Means Clustering*. Penentuan k optimal menggunakan *Elbow Method* dan *Silhouette Analysis*. Validitas klaster diukur dengan *Silhouette Score*, *Davies–Bouldin Index*, dan *Calinski–Harabasz Index*.
4. Hasil akhir berupa profil karakteristik setiap klaster dan rekomendasi strategi pemasaran digital. Rekomendasi akan divalidasi melalui *expert judgement* (diskusi terstruktur) dengan pemilik usaha atau tim pemasaran. Penelitian tidak mencakup pengembangan sistem *real-time*, integrasi API, pengujian A/B, atau analisis sentimen teks komentar.

5. Setiap observasi harian diperlakukan independen (bukan analisis *time-series*). Nilai negatif atau *missing value* ditangani pada tahap *preprocessing*. Data telah diverifikasi internal dan mencerminkan performa riil.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dirumuskan, tujuan penelitian ini adalah:

1. Menerapkan tahapan *preprocessing* data (pembersihan, penanganan nilai anomali, *feature engineering*, dan normalisasi) pada *dataset* agregat harian performa akun TikTok.
2. Menentukan jumlah klaster (k) optimal yang paling merepresentasikan pola sebaran performa harian melalui evaluasi *Elbow Method* dan *Silhouette Score*.
3. Menganalisis dan mendeskripsikan karakteristik serta profil masing-masing klaster hasil segmentasi berdasarkan rata-rata dan distribusi variabel metrik.
4. Merumuskan dan memvalidasi rekomendasi strategi optimasi pemasaran digital yang terukur dan kontekstual bagi Perusahaan Kecap Djoe-Hoa melalui diskusi dengan pemangku kepentingan.

1.5 Manfaat Penelitian

1.5.1 Manfaat Teoritis

- a. Menambah referensi keilmuan di bidang Sistem Informasi dan *Data Mining* mengenai penerapan *unsupervised learning* (khususnya *K-Means Clustering*) untuk segmentasi performa konten pada platform media sosial video pendek.

- b. Menyediakan kerangka metodologis yang dapat diadaptasi atau dikembangkan oleh penelitian selanjutnya, misalnya dengan membandingkan algoritma klusterisasi lain (*DBSCAN*, *Hierarchical*, *Gaussian Mixture*) atau mengintegrasikan teknik prediktif berbasis *supervised learning*.

1.5.2 Manfaat Praktis

- a. Bagi Perusahaan Kecap Djoe-Hoa
Memberikan dasar pengambilan keputusan yang objektif dalam perencanaan konten, penjadwalan unggahan, alokasi anggaran promosi, serta pengembangan format video yang disesuaikan dengan segmen performa konten yang terbukti efektif.
- b. Bagi Praktisi Pemasaran Digital & UMKM
Menunjukkan bagaimana data analitik organik TikTok dapat diolah secara sistematis menjadi *actionable insight* untuk meningkatkan *Return on Content Investment* (ROCI) tanpa bergantung pada asumsi subjektif.
- c. Bagi Peneliti Selanjutnya
Menyediakan *pipeline* analisis data, dokumentasi parameter *preprocessing*, serta hasil evaluasi klaster yang dapat direplikasi, divalidasi, atau diperluas cakupannya pada platform lain atau *dataset* yang lebih besar.