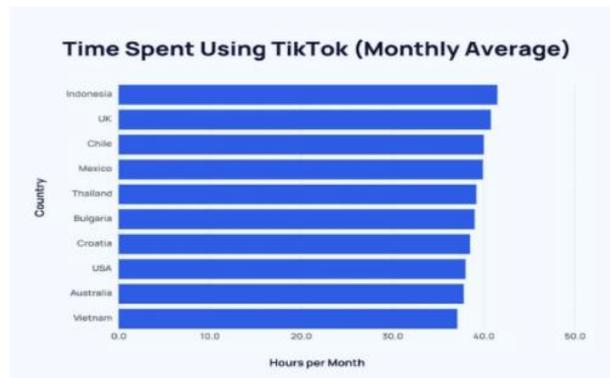


I. PENDAHULUAN

1.1 Latar Belakang

Penggunaan media sosial telah menjadi bagian integral dari kehidupan masyarakat global. Menurut laporan dari Data Reportal (Kemp, 2024), jumlah pengguna media sosial di seluruh dunia mencapai 5,04 miliar pada awal tahun 2024, atau sekitar 62,3% dari populasi global. Angka ini menunjukkan pertumbuhan yang konsisten setiap tahunnya, seiring dengan peningkatan penetrasi internet dan penggunaan smartphone secara global. Fenomena global ini juga mendorong perubahan signifikan dalam pola komunikasi dan konsumsi informasi, terutama di kalangan generasi muda. Salah satu platform yang mencerminkan perubahan tersebut adalah TikTok, sebuah aplikasi berbasis video pendek yang berhasil menarik perhatian masyarakat di berbagai negara. Hingga tahun 2023, TikTok telah mencapai lebih dari 1 miliar pengguna aktif bulanan secara global, menjadikannya sebagai salah satu platform dengan pertumbuhan tercepat sepanjang sejarah media sosial (Oberlo, 2024).

Pertumbuhan TikTok yang sangat pesat tidak hanya mencerminkan popularitasnya, tetapi juga menunjukkan perubahan pola perilaku digital pengguna yang semakin aktif, tidak hanya sebagai konsumen pasif tetapi juga sebagai kreator konten untuk tujuan hiburan, edukasi, maupun komersial. Platform ini mendorong kreativitas visual dan interaksi yang tinggi, sehingga menciptakan pola penggunaan yang kompleks dan beragam (Zannettou et al., 2024). Pada September 2024, TikTok mencatat sekitar 2,5 miliar kunjungan bulanan, dengan lebih dari 70% akses melalui perangkat seluler. Rata-rata pengguna melihat 4,46 halaman per kunjungan dan menghabiskan waktu sekitar 8 menit 18 detik di platform tersebut (Duarte, 2024). Secara global, Indonesia tercatat sebagai negara dengan waktu penggunaan TikTok tertinggi, yakni mencapai 41 jam 35 menit per bulan (Duarte, 2024).



Gambar 1. Negara-negara teratas pengguna TikTok
Sumber : www.explodingtopics-com

Meskipun TikTok telah menjadi fenomena global dengan lebih dari 1 miliar pengguna aktif bulanan di seluruh dunia (Statista, 2024), studi akademik yang secara khusus mengeksplorasi segmentasi pengguna berdasarkan perilaku interaksi masih tergolong terbatas untuk media sosial. Sebagian besar penelitian sebelumnya cenderung fokus pada konten viral, strategi pemasaran, persepsi pengguna, atau algoritma rekomendasi, bukan pada pemetaan struktur perilaku pengguna secara kuantitatif dan sistematis (Hindarto, 2022). Padahal, pendekatan berbasis data sangat diperlukan untuk memahami karakteristik pengguna secara lebih akurat. Informasi semacam ini penting bagi pengembang aplikasi, pemasar digital, maupun peneliti sosial, karena dapat menjadi dasar untuk personalisasi konten, segmentasi pasar, dan pengembangan sistem rekomendasi yang lebih adaptif terhadap perilaku pengguna (Han et al., 2024).

Data mining merupakan proses menggali dan menganalisis data dalam jumlah besar dari basis data untuk menemukan pola, kecenderungan, atau keterkaitan tersembunyi yang memiliki nilai guna. Proses ini memanfaatkan beragam pendekatan, seperti metode statistik, teknik matematika, dan algoritma *machine learning*, guna menghasilkan pengetahuan yang mendukung proses pengambilan keputusan (Wahono, 2023). Salah satu pendekatan dalam *data mining* adalah *unsupervised learning*, yaitu metode pembelajaran mesin tanpa label target yang bertujuan menemukan pola tersembunyi dalam data tanpa arahan eksplisit (Nurhalizah et al., 2024). Pendekatan ini sangat relevan dalam konteks analisis pengguna TikTok, karena data interaksi yang digunakan tidak memiliki label klasifikasi tertentu, melainkan berupa data seperti jumlah tampilan, suka, komentar, unduhan, dan pembagian video. Berdasarkan pertimbangan tersebut, penelitian ini menerapkan dua algoritma utama dalam *data mining*, yaitu *Partitioning Around Medoids* (PAM) untuk melakukan klasterisasi pengguna, dan *Frequent Pattern Growth* (FP-Growth) untuk menemukan pola asosiasi antaraktivitas pengguna dalam *platform* TikTok.

Algoritma *Partitioning Around Medoids* (PAM) merupakan salah satu metode klastering berbasis partisi yang menggunakan objek aktual dalam dataset sebagai pusat klaster, yang disebut *medoid*. *Medoid* dipilih karena mewakili objek paling sentral dalam suatu klaster, berbeda dengan algoritma K-Means yang menggunakan nilai rata-rata (*centroid*) sebagai pusat klaster. Keunggulan utama PAM adalah kemampuannya dalam menangani *outlier* dan *noise*, karena pemilihan *medoid* tidak mudah terpengaruh oleh nilai ekstrem (Bipul Hossen & Rabiul Auwul, 2020). PAM membagi data ke dalam k klaster berdasarkan jarak terkecil antara setiap data dan *medoid*-nya, yang dihitung menggunakan metrik jarak tertentu, seperti *Euclidean distance*. Metode ini sangat cocok digunakan

untuk data interaksi pengguna media sosial yang bersifat numerik dan memiliki persebaran nilai yang luas, karena mampu menangkap representasi yang lebih stabil dari kelompok data (Aeni et al., 2020). Penelitian sebelumnya oleh Seimahaira (2022) dalam judul “Analisa Pola Pengguna Youtube Trending Menggunakan Algoritma *Partitioning Around Medoids* dan *Frequent Pattern Growth*” juga membuktikan bahwa PAM efektif dalam mengelompokkan pengguna berdasarkan intensitas interaksi, seperti jumlah *view*, *like*, *share*, dan *comment* pada platform YouTube. Hasil klaster dari metode ini mampu mengungkap kategori pengguna, serta memetakan perilaku digital secara lebih terstruktur.

Selain *clustering*, pemahaman terhadap keterkaitan antaraktivitas pengguna juga sangat penting untuk mengungkap pola perilaku secara menyeluruh. Oleh karena itu, penelitian ini juga menggunakan algoritma *Frequent Pattern Growth* (FP-Growth), yaitu metode pencarian pola asosiasi dalam dataset transaksi yang memiliki keunggulan dalam efisiensi dan kecepatan dibandingkan algoritma *Apriori*. FP-Growth membentuk struktur FP-Tree untuk menyimpan data transaksi secara kompak, sehingga tidak perlu eksplorasi berulang pada dataset (Paila et al., 2025). Dengan pendekatan ini, dapat ditemukan aturan asosiasi antarinteraksi yang sering dilakukan secara bersamaan, seperti hubungan antara pengguna yang sering menonton video dengan kecenderungan mereka untuk memberikan komentar atau membagikan konten. Studi oleh Munanda dan Monalisa (2021) menyatakan bahwa FP-Growth mampu menghasilkan pola asosiasi yang signifikan dengan waktu komputasi yang lebih cepat dibandingkan *Apriori*, terutama pada dataset berskala besar. Dalam penelitian Seimahaira (2022), FP-Growth juga digunakan untuk mengidentifikasi keterkaitan perilaku pengguna YouTube dan berhasil menemukan kombinasi aktivitas yang sering muncul bersama, seperti *like* dan *comment* atau *share* dan *view*. Hal ini memperkuat bahwa metode FP-Growth layak diterapkan dalam analisis perilaku pengguna TikTok yang memiliki karakteristik data serupa.

Beberapa penelitian telah menerapkan algoritma *Partitioning Around Medoids* (PAM) dan *Frequent Pattern Growth* (FP-Growth) untuk menganalisis dan menemukan asosiasi atau pola dalam sebuah data. Contohnya penelitian yang dilakukan oleh (Seimahaira, 2022) dengan judul Analisa Pola Pengguna Youtube Trending Menggunakan Algoritma *Partitioning Around Medoids* dan *Frequent Pattern Growth*, penelitian (Auliasari & Mariza Kertaningtyas, 2022) dengan judul Analisis *Cluster* Atribut Audio pada Lagu Terpopuler Aplikasi TikTok, penelitian (Musu et al., 2024) dengan judul *Enhancing TikTok Account Performance with Data*

Pattern Identification. Penelitian ini menggunakan algoritma *Partitioning Around Medoids* (PAM) dan *Frequent Pattern Growth* (FP-Growth) untuk memperoleh pemahaman yang lebih komprehensif mengenai pola perilaku pengguna TikTok secara global. Melalui pendekatan ini, penelitian tidak hanya memetakan segmentasi pengguna berdasarkan intensitas interaksi seperti melihat, menyukai, mengunduh, membagikan, dan mengomentari video, tetapi juga mengungkap keterkaitan antaraktivitas yang sering muncul bersamaan. Kedua metode ini memberikan perspektif dua arah, dari struktur kelompok pengguna melalui hasil klasterisasi, hingga identifikasi hubungan asosiasi antarinteraksi berdasarkan frekuensi kemunculannya.

Dengan demikian, hasil analisis ini dapat menjadi dasar awal bagi pengambilan keputusan dalam pengelompokan pengguna serta pemahaman pola interaksi secara umum, yang berpotensi digunakan dalam segmentasi pasar dan pengembangan strategi *engagement* pengguna secara lebih terarah. Meskipun penelitian ini belum mencakup konteks konten atau demografi pengguna secara spesifik, hasil yang diperoleh tetap memberikan gambaran awal mengenai perilaku pengguna TikTok secara kuantitatif, yang dapat dijadikan pijakan untuk studi lanjutan dan perumusan strategi digital berbasis data.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana pengelompokan (*clustering*) pengguna TikTok berdasarkan pola interaksinya?
2. Pola asosiasi apa saja yang dapat ditemukan dari aktivitas pengguna TikTok?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, tujuan dari penelitian ini adalah:

1. Mengelompokkan pengguna TikTok berdasarkan intensitas aktivitasnya menggunakan algoritma *Partitioning Around Medoids* (PAM).
2. Mengidentifikasi pola asosiasi antaraktivitas pengguna TikTok menggunakan algoritma *Frequent Pattern Growth* (FP-Growth).

1.4 Manfaat Penelitian

Penelitian ini memiliki beberapa manfaat yang diharapkan dapat berguna bagi berbagai pihak, antara lain:

1. Manfaat Teoritis : Menambah wawasan mengenai penggunaan algoritma data mining seperti PAM dan FP-Growth dalam menganalisis pola penggunaan media sosial.
2. Manfaat praktis : Memahami penggunaan media sosial dan menemukan pola yang mungkin tidak disadari.
3. Pengembang Aplikasi Media Sosial: Memberikan data terkait preferensi dan pola penggunaan aplikasi TikTok.

1.5 Batasan Masalah

Untuk menjaga fokus dan efektivitas penelitian, terdapat beberapa batasan masalah yang ditetapkan dalam penelitian ini, yaitu:

1. Data yang digunakan dalam penelitian ini adalah dataset pengguna TikTok yang bersumber dari Kaggle. Dataset ini diunggah oleh Ramin Husein, seorang pakar data dari Plymouth, Inggris, Britania Raya, serta mahasiswa Magister Operasi dan Manajemen Rantai Pasokan di *University of Plymouth* pada tahun 2024. Data yang tersedia dalam dataset ini mencakup 19.382 baris data dan 12 kolom atribut.
2. Atribut kategori yang digunakan untuk *clustering* dan asosiasi adalah aktivitas pengguna, seperti jumlah tampilan video (*video_view_count*), jumlah suka (*video_like_count*), jumlah berbagi (*video_share_count*), jumlah unduhan (*video_download_count*), dan jumlah komentar (*video_comment_count*). Atribut yang digunakan dalam penelitian ini dipilih berdasarkan aktivitas utama pengguna TikTok, yang mencerminkan interaksi mereka dengan konten.
3. Penelitian ini bersifat eksploratif dengan tujuan utama untuk menggali dan memahami pola interaksi pengguna TikTok berdasarkan metrik kuantitatif (*view, like, comment, download, dan share*). Penelitian ini tidak bertujuan untuk membangun model prediksi atau klasifikasi, melainkan fokus pada identifikasi pola dan klusterisasi melalui pendekatan data mining seperti *clustering* dan *asosiasi*.
4. Analisis clustering menggunakan algoritma *Partitioning Around Medoids* (PAM), dan tidak dibandingkan dengan algoritma *clustering* lainnya seperti K-Means atau DBSCAN. Analisis asosiasi pola menggunakan algoritma FP-Growth tanpa perbandingan dengan metode lain seperti Apriori.
5. Perhitungan jarak untuk klusterisasi hanya menggunakan metode *Euclidean Distance*, dan evaluasi klusterisasi hanya menggunakan matrik *Davies-Boulden Index* (DBI). Pada Penelitian ini tidak menganalisis faktor demografis pengguna seperti usia, jenis kelamin, atau lokasi geografis.