

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi telah membawa perubahan signifikan dalam masyarakat. Salah satu aspek yang mengalami perkembangan pesat adalah media sosial. Media sosial merupakan kumpulan aplikasi berbasis internet yang memungkinkan terjadinya berbagai jenis interaksi yang sebelumnya tidak dapat diakses oleh masyarakat umum. Melalui media sosial, pengguna dapat mempresentasikan diri, berinteraksi, bekerja sama, berbagi, dan berkomunikasi dengan pengguna lainnya, membentuk ikatan sosial secara virtual (Prihatiningsih, 2017). Salah satu platform media sosial dengan angka pengguna yang banyak adalah *instagram*

Instagram adalah platform media sosial yang memiliki lebih dari 61.610.000 pengguna aktif di Indonesia adalah *instagram*. Dengan jumlah pengguna yang signifikan, *instagram* telah menjadi salah satu media sosial yang paling populer dan berpengaruh dalam kehidupan sehari-hari (Sari & Basit, 2020). *Instagram* menyediakan layanan berbagi foto ataupun video secara *online*, banyak digunakan oleh para remaja. Mereka dapat mengunggah foto dan video ke dalam galeri, yang kemudian dapat diedit dengan berbagai filter dan diiringi dengan *caption*.

Unggah tersebut dapat dibagikan secara publik atau hanya dapat dilihat oleh pengikut. Selain digunakan untuk berkomunikasi, media sosial juga berperan sebagai media hiburan penyampaian berita, dan sumber informasi yang disematkan ke dalam *caption* pada postingan. Beberapa pengguna *instagram* memanfaatkan akun mereka untuk membagikan informasi tentang berita dan fenomena kehidupan sehari – hari. Setiap konten atau postingan yang diunggah tentunya memiliki makna yang disukai oleh pengguna tersebut. Dengan meningkatnya jumlah pengguna yang melakukan posting di *Instagram*, semakin

mudah bagi kita untuk mengenali ketertarikan dari pengguna tersebut melalui *caption* pada postingan.

Dengan banyaknya postingan pada profil *instagram* hal tersebut dapat menunjukkan minat khusus dari pengguna media sosial terkait topik tertentu. Pengguna media sosial yang memiliki banyak pengikut juga memiliki dampak signifikan. Namun banyaknya postingan dari setiap pengguna media sosial juga menciptakan masalah baru, dimana pengguna lain harus membaca kata per-kata pada *caption* di setiap postingan untuk mengetahui ketertarikan dari profil tersebut sebagaimana pada platform *instagram* tidak memiliki fitur untuk mengkategorikan ketertarikan dari setiap profil *instagram*.

Pengenalan dan analisis data merupakan elemen penting dalam berbagai bidang, termasuk ilmu komputer dan statistika. Dalam konteks ini, *unsupervised learning* adalah salah satu pendekatan yang digunakan untuk menggali informasi dari data yang tidak memiliki label atau anotasi sebelumnya. *Unsupervised learning* memungkinkan sistem komputer untuk secara otomatis mengenali pola, struktur, atau keterkaitan yang tersembunyi dalam data.

Terdapat penelitian terdahulu yang telah dilakukan oleh Binuri Ayu Dwiarni dan Budi Setiyono yang berjudul “*Akusisi dan Clustering Data Sosial Media Menggunakan Algoritma K-Means sebagai Dasar untuk Mengetahui Profil Pengguna*” pada tahun 2019 (Dwiarni & Setiyono, 2019). Dari hasil validasi penelitian ini adalah hasil perhitungan *clustering* secara manual dibanding dengan penghitungan *clustering* mirip meskipun tidak 100% sama. Yang perlu diperhatikan bahwa algoritma *K-means* masuk ke dalam *unsupervised* dan dapat menghasilkan hasil yang berbeda karena adanya proses randomisasi pada algoritma *K-Means*. Proses studi literatur dalam mencari referensi tentang metode *clustering* beberapa terdapat referensi yang menggunakan algoritma *K-Means*, dalam sistem ini, akan diterapkan metode *Clustering K-Means* untuk memproses data teks dan mengelompokkan ke dalam kategori yang sesuai. Metode ini akan memungkinkan identifikasi pola dan kesamaan dalam penggunaan kata-kata pada

caption instagram. Dengan demikian, setiap *caption* akan ditempatkan dalam *cluster* yang mencerminkan minat dari profil tersebut.

Dalam penelitian ini, akan dikembangkan sebuah sistem yang menggunakan metode *clustering* untuk dapat menghasilkan kategori dari setiap profil *instagram*. Sistem akan menggunakan data teks yang berasal dari *caption* pada setiap postingan profil *instagram*. Data tersebut akan dikumpulkan sebanyak 1000 data dari referensi terdahulu dengan judul “*Clustering K-Means* Jenis Kata Pada Laporan Kegiatan Kuliah Kerja Nyata Universitas Lampung” yang menggunakan 1000 data dan menghasilkan 2 *cluster* atau kategori (Ademariana et al., 2021). Pada rentan waktu Agustus hingga Oktober 2022, diambil sebanyak 1000 data dari *caption* postingan di *instagram*.

Penelitian ini mengambil data dari berbagai sumber yang kemudian dikategorikan menjadi empat profil berdasarkan analisis individu untuk mengetahui ketertarikan pada setiap profil. Profil pertama, "nexcarnos", didominasi oleh pembahasan seputar makanan, sementara profil kedua, "ariefmuhammad", cenderung lebih fokus pada perjalanan dan wisata, atau lebih dikenal sebagai travel. Selanjutnya, profil "andreanadianbimo" lebih cenderung mengungkapkan minat dalam pembahasan olahraga, sementara profil "tasyafarasya" secara dominan tertarik pada topik kecantikan. Dalam penelitian ini, setiap profil akan diwakili dengan 250 data, sehingga proporsinya menjadi seimbang dalam analisis yang dilakukan.

Maka dari itu dalam penelitian ini akan dikembangkan sebuah sistem yang diharapkan dapat membantu dalam menentukan kesesuaian antara kriteria dengan profil *instagram* serta memberikan informasi tanpa perlu melakukan analisis manual. Sistem ini akan membantu dalam memberikan informasi mengidentifikasi ketertarikan profil berdasarkan minat setiap profil *instagram*.

1.2 Rumusan Masalah

Dari paparan latar belakang yang telah dipaparkan di atas, muncul sebuah rumusan masalah yang menjadi dasar dari terlaksananya penelitian ini adalah,

Bagaimana hasil implementasi metode *clustering K-Means* untuk mengetahui ketertarikan profil pengguna media sosial media berdasarkan *caption* profil media sosial pada platform *instagram* ?

1.3 Tujuan

Berdasarkan rumusan masalah yang telah di paparkan diatas, penelitian ini dilakukan dengan tujuan sebagai berikut :

1. Mengetahui hasil penerapan metode *Clustering K-Means* untuk mengetahui ketertarikan profil *instagram* berdasarkan text *caption* postingan.
2. Mempersingkat dalam mengetahui ketertarikan setiap profil *instagram* dari *caption* pada setiap postingan.
3. Mendapatkan data valid dari ketertarikan profil *instagram*.

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah :

1. Bagi Mahasiswa

Penelitian ini diharapkan bisa memberikan manfaat bagi mahasiswa yang nantinya berkesempatan untuk menggunakan penelitian ini sebagai acuan guna membantu proses penelitian berikutnya.

2. Bagi Masyarakat

Memberikan kepada masyarakat tentang informasi ketertarikan setiap profil *instagram* tanpa perlu membaca secara manual pada setiap postingan.