

## I. PENDAHULUAN

### 1.1. Latar Belakang

Dalam era digital saat ini, perkembangan teknologi informasi dan komunikasi telah mengubah banyak aspek kehidupan manusia termasuk media untuk mereka berinteraksi dan berkomunikasi menjadi sangat beragam. Adanya kehadiran media sosial menjadikan antara ruang privasi seseorang dan publik dapat melebur begitu cepat. Saat ini kehidupan masyarakat tidak terlepas dari media sosial, yang mana media sosial *online* bermanfaat sebagai sarana interaksi sosial hanya dengan menggunakan internet (Norzhela & Jaya, 2022). Dengan menggunakan media sosial, para penggunanya dapat saling melihat informasi, menjadikan tempat arsip, berkomunikasi, berinteraksi, melakukan simulasi sosial, melihat dan mengupload konten yang relevan, hingga memposting aktivitas masing-masing (Setiadi, 2016).

Indonesia dengan populasi yang besar dan tingkat adopsi teknologi yang semakin meningkat, telah melihat beragam platform media sosial mendapatkan popularitas. Tercatat hingga bulan Januari 2023, terdapat sebanyak 167,0 juta pengguna media sosial di Indonesia, yaitu setara dengan 60,4% dari total populasi di dalam negeri (Kemp, 2023). Jumlah pengguna aktif media sosial pada Januari 2023 mengalami penurunan 12,57% dibandingkan pada tahun sebelumnya yaitu sebanyak 191 juta jiwa. Penurunan itu pun menjadi yang pertama kali terjadi dalam satu dekade terakhir (Widi, 2023).

Menurut Hasya (2023), terdapat beberapa media sosial yang populer di Indonesia, antara lain sebagai berikut.

**Tabel 1.** Pengguna Media Sosial

No	Media Sosial	Jumlah Pengguna
1	WhatsApp	92,1%
2	Instagram	86,5%
3	Facebook	83,8%
4	TikTok	70,8%
5	Telegram	64,3%
6	Twitter	60,2%
7.	Facebook Messenger	51,9%

Sumber: (Data Olahan, 2023)

Berdasarkan Tabel 1 diketahui bahwa jumlah pengguna media sosial di Indonesia pada tahun 2022 hingga Januari 2023 tertinggi yaitu Whatsapp dengan jumlah yang mencapai 92,1%. Aplikasi WhatsApp atau yang biasa di singkat dengan WA, merupakan teknologi *Instant Messaging* yang dapat mengirim pesan

seperti *Short Message Service* atau layanan pemberian pesan singkat, melakukan panggilan dan *video call*, dan juga dapat mengirimkan berbagai macam file seperti gambar, video, suara dan dokumen baik secara pribadi maupun grup (Harianto et al., 2022).

Penggunaan WhatsApp sebagai media sosial hampir sering digunakan manusia dalam kehidupan sehari-hari. Hal ini membuat komunikasi dan interaksi sosial melalui pesan, gambar, audio, video, dokumen, dan hal yang unik lainnya termasuk diskusi menjadi lebih mudah (Koten et al., 2022). Aplikasi ini memiliki fitur pendukung yang dapat digunakan oleh pengguna untuk berbagi informasi dan berbagai konten melalui WhatsApp *story*.

Whatsapp telah berkembang menjadi media pilihan karena kemampuannya untuk beradaptasi dengan berbagai kondisi sinyal. Walau sinyal lemah, pengiriman data teks, suara, foto, audio, dan video tanpa waktu tunggu masih dapat dilakukan. Penggunaan aplikasi WhatsApp juga terbilang menarik dikarenakan tidak ada iklan yang mengganggu (Rahartri, 2019). Selain itu, WhatsApp juga menyediakan fitur obrolan grup, di mana beberapa pengguna dapat berpartisipasi dalam obrolan yang sama dalam satu percakapan. Fitur ini membantu dan memudahkan pengguna untuk berkomunikasi secara bersamaan dengan orang lain tanpa perlu berinteraksi secara langsung. Bahkan, fitur ini dapat dijadikan sebagai wadah untuk mendiskusikan topik penting serta memfasilitasi pengambilan keputusan (Shidiq & Warnilah, 2019).

Aplikasi Whatsapp dengan keberadaannya yang telah lama dan semakin menjadi prioritas terhadap layanan media sosial, memerlukan identifikasi mengenai kepuasan penggunaannya. Pengukuran kepuasan ini didasarkan pada motif penggunaan media yang dikemukakan oleh Teori Mc-Quail yaitu motif informasi (*Surveillance*), motif identitas pribadi (*Personal Identity*), motif integrasi dan interaksi sosial (*Personal Relationship*), dan motif hiburan (*Diversion*) (Hilabi & Priati, 2018). Selain itu, dalam upaya untuk mengetahui tingkat kepuasan pengguna, peneliti menggunakan metode *end user computing satisfaction* (EUCS) sebagai referensi dalam pengumpulan data. Instrumen *end user computing satisfaction* (EUCS) yang mencakup lima komponen yaitu isi (*content*), keakuratan (*accuracy*), bentuk (*format*), kemudahan penggunaan (*ease of use*) dan ketepatan waktu (*timeliness*) juga digunakan sebagai acuan variabel (Saputri & Alvin, 2020). Pada hasil penelitian Setyawan & Firmanto (2022) juga dijelaskan bahwa variabel pada metode *Delone and McLean*, yang disempurnakan menjadi *Delone and Mclean Models* yaitu *information quality*, *system quality*, dan *service quality* akan berpengaruh positif pada *use & user satisfaction* dan *net benefit* (hasil akhir) terhadap kepuasan penggunaan sistem.

Dalam rangka mengukur kepuasan pengguna aplikasi WhatsApp, penelitian ini akan menerapkan sejumlah metode dan variabel yang relevan. Pengolahan data akan melibatkan variabel-variabel yang berdasarkan dari Teori Mc-Quail yaitu motif integrasi dan interaksi sosial (*Personal Relationship*), variabel dari *End User Computing Satisfaction* (EUCS) seperti kemudahan penggunaan (*ease of use*). Sementara itu, variabel dari model Delone and McLean akan mengevaluasi kualitas informasi, kualitas sistem secara keseluruhan, dan kualitas layanan yang diberikan oleh WhatsApp. Dari sejumlah instrumen tersebut dapat diolah pada pemanfaatan data *mining*.

Pemanfaatan data *mining* dalam penelitian ini memiliki peran yang krusial, yaitu untuk mengeksplorasi dan menggali pola-pola yang tidak terlihat secara langsung dalam data kepuasan pengguna WhatsApp. Analisis data mining memberikan kemampuan untuk mengidentifikasi korelasi, tren, dan pola-pola kompleks yang sulit atau bahkan tidak dapat ditemukan secara manual (Handoko et al., 2020). Dengan menggunakan teknik data *mining*, penelitian ini dapat mengurai set data yang besar dan kompleks, menjelajahi hubungan antar variabel, dan menyoroti faktor-faktor yang memiliki dampak signifikan terhadap kepuasan pengguna WhatsApp.

Teknik data *mining* yang paling sering digunakan adalah *clustering* dan *classification* (Patimah et al., 2021). Dalam penelitian ini, teknik data *mining* yang digunakan adalah *clustering*. *Clustering* adalah suatu teknik yang digunakan untuk menganalisis pengelompokan terhadap data, di mana data akan dibagi menjadi kelompok-kelompok berdasarkan kemiripannya (Nofriansyah & Nurcahyo, 2017). Ada dua metode dalam *clustering* atau yang dikenal juga sebagai analisis *cluster* yaitu metode hirarki dan non hirarki. Metode hirarki merupakan metode yang membuat sebuah dekomposisi berhirarki (tingkatan) dari himpunan data berdasarkan kesamaan karakteristik objeknya. Sedangkan metode non hirarki digunakan untuk pengelompokan objek, dimana jumlah *cluster* yang akan dibentuk dapat ditentukan sebelumnya. Terdapat beberapa jenis analisis *cluster* yang menggunakan metode hirarki, antara lain *single linkage*, *complete linkage*, *average linkage*, *centroid*, *ward* dan *median cluster*. Sedangkan dalam analisis *cluster* non hirarki, antara lain adalah metode *K-Means*, *DBSCAN*, dan *Gaussian Mixture Modelling* (GMM) (Widyadhana et al., 2021).

Beberapa penelitian yang terkait dengan *clustering* sangatlah banyak, penelitian yang telah dilakukan oleh Patimah et al. (2021) menunjukkan bahwa pengelompokan kepuasan pelanggan Shopee dilakukan dengan menggunakan algoritma *K-Means*. *K-Means* merupakan salah satu algoritma *clustering*, di mana

*K-Means* akan menghasilkan kelompok berdasarkan kepada kemiripannya, sehingga metode ini cocok untuk digunakan dalam penelitian ini. *Cluster* yang digunakan memiliki jumlah 2, 3, 4, dan 5. *Cluster-cluster* ini dievaluasi menggunakan *Davies Bouldin Index* (DBI), di mana *cluster* dengan nilai DBI terkecil dianggap paling optimal. Hasil penelitian menunjukkan bahwa *K-Means* dengan  $k=2$  menghasilkan *cluster* paling optimal, dengan nilai DBI sebesar 1.587617820812729.

Hasil pada Penelitian Pradana et al. (2021) yang juga menggunakan algoritma *K-Means* menunjukkan bahwa algoritma tersebut efektif untuk mengelompokkan data. Data dari penelitian ini terbagi menjadi 2 *cluster* yaitu *cluster* 0 (puas) dan *cluster* 1 (tidak puas ataupun netral). Pengelompokan *cluster* tersebut berdasarkan *dataset* yang dimiliki, dimana responden mengisi data dan menghasilkan 2 jenis *cluster* tersebut. Adapun hasil dari proses *clustering* adalah sebanyak 1303 data masuk kategori *cluster* 0 atau sebesar 65% dan 697 data masuk kategori *cluster* 1 atau sebesar 35%.

Berbeda dengan penelitian yang telah dijelaskan sebelumnya yang fokus pada penggunaan algoritma *K-means*, pada hasil penelitian Riyana & Dwidayanti (2021) menunjukkan bahwa metode analisis data yang digunakan yaitu dengan melakukan perbandingan antara algoritma *K-Means* dan *Fuzzy C-Means* dimana uji validitas *cluster* menggunakan *Dunn-Index* dan *Davies Bouldin-Index* akan memperoleh hasil *cluster* yang optimal. Hasil penelitian menunjukkan bahwa pengujian *K-Means clustering* menghasilkan nilai akurasi yang lebih besar yaitu 1,165219 dibandingkan *Fuzzy C-Means* yang memperoleh nilai sebesar 1,74298. Sehingga *K-Means clustering* diambil untuk menentukan pengelompokan daerah penyebaran Covid-19 per provinsi Indonesia. Selanjutnya pada penelitian tentang perbandingan *K-Means* dan *K-Medoids*, Pada Data Potensi Kebakaran Hutan dan Lahan oleh Athifaturrofifah et al. (2019) menunjukkan hasil bahwa *K-Means* dapat memberikan hasil pengelompokan yang lebih baik daripada *K-Medoids* dikarenakan nilai *Silhouette Coefficient* metode *K-Means* mencapai 0,558, sedangkan Nilai *Silhouette Coefficient* metode *K-Medoids* sejumlah 0,529.

Pada penelitian lainnya yang telah diteliti oleh Sapriyanti & Rianto (2020) menunjukkan bahwa pada saat proses pengelompokan menggunakan metode *K-Means* dan metode *Single Linkage*, terdapat perbedaan dalam hal jumlah *Agent* dalam sebuah *cluster*, anggota pada masing masing *cluster*, dan waktu eksekusi untuk membangun sebuah *cluster*. Dalam hal waktu eksekusi untuk membangun sebuah *cluster* dengan 20 *agent contact center*, metode *K-Means* lebih unggul dibanding dengan metode *Single Linkage*. Dengan mencari nilai *centroid distance*, didapatkan bahwa *cluster* yang terbentuk dari metode *K-Means* lebih ideal

dibanding metode *Single Linkage*. Dimana jumlah *member* pada *cluster* dengan metode *K-Means* yaitu C0 (3), C2 (9) dan C3 (8) sedangkan dengan metode *Single Linkage* yaitu C0 (8), C2 (9) dan C3 (3). Kemudian pada penelitian Taslim et al. (2021) yang dimana terdapat dua pemodelan yaitu model yang pertama adalah *clustering* hanya menggunakan *K-means*, yang kedua adalah *K-means* yang dioptimasi dengan algoritma genetika. Data yang digunakan melibatkan sembilan kabupaten dan kota serta 282 dusun dan desa di Provinsi Riau baik itu desa baru maupun desa lama. Adapun data yang digunakan adalah data target pemanfaatan air bersih PAMSIMAS tahun 2019. Hasil yang didapatkan yaitu klasterisasi tanpa optimasi, menghasilkan DBI sebesar 2.164763 sedangkan hasil klasterisasi dengan optimasi genetika pada nilai *centroid* menghasilkan nilai DBI sebesar 2.06894. Dari hasil tersebut terlihat bahwa melakukan optimasi pada nilai *centroid* menghasilkan *cluster* yang lebih baik dibandingkan dengan klasterisasi *K-means* tanpa optimasi *centroid*.

Dari penelitian-penelitian *clustering* terkait yang telah dipaparkan di atas, dapat disimpulkan bahwa algoritma *K-Means* adalah algoritma yang baik. Metode *K-Means* merupakan metode *clustering* yang paling sederhana dan umum. Hal ini dikarenakan *K-Means* mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien (Arifin et al., 2017).

Berdasarkan apa yang telah diuraikan di atas, penelitian mengenai analisis *cluster* kepuasan pengguna terhadap aplikasi Whatsapp menggunakan algoritma *K-means* menjadi hal yang sangat penting. Penelitian ini dilakukan untuk mengelompokkan data kepuasan pengguna dalam menggunakan aplikasi WhatsApp di Indonesia menggunakan metode *K-Means*.

### **1.2. Rumusan Masalah**

Berdasarkan latar belakang permasalahan yang telah dipaparkan, maka pokok permasalahan pada penelitian ini adalah :

1. Bagaimana hasil implementasi algoritma *K-Means* untuk mengelompokkan pengguna aplikasi WhatsApp di Indonesia?
2. Bagaimana cara menentukan jumlah *cluster* yang optimal untuk mengetahui kepuasan pengguna aplikasi WhatsApp di Indonesia?

### **1.3. Batasan Masalah**

Dalam penelitian ini, terdapat beberapa batasan masalah yang ditentukan agar fokus penelitian dapat terjaga dan hasil yang didapatkan lebih spesifik. Batasan-batasan tersebut antara lain:

1. Penelitian ini akan fokus pada data pengguna WhatsApp di Indonesia. Data dari platform media sosial negara lain tidak akan dimasukkan dalam analisis ini.
2. Meskipun ada berbagai metode *clustering* yang bisa digunakan, penelitian ini akan memfokuskan analisis pada algoritma *K-Means*.
3. Penelitian ini akan dilakukan dalam kurun waktu tertentu, sehingga data yang didapatkan hanya akan mencerminkan keadaan pada periode tersebut dan akan berbeda dengan periode lainnya.

#### **1.4. Tujuan Penelitian**

Berdasarkan rumusan masalah yang telah didefinisikan sebelumnya, tujuan penelitian ini dirumuskan sebagai berikut:

1. Mengetahui hasil implementasi algoritma *K-Means* untuk mengelompokkan pengguna aplikasi WhatsApp di Indonesia.
2. Mengetahui cara menentukan jumlah *cluster* yang optimal untuk mengetahui kepuasan pengguna aplikasi WhatsApp di Indonesia.

#### **1.5. Manfaat Penelitian**

Berdasarkan tujuan yang telah ditetapkan pada penelitian ini, beberapa manfaat yang diharapkan antara lain:

1. Penelitian ini dapat memberikan kontribusi terhadap literatur ilmiah dalam bidang analisis *cluster* kepuasan pengguna terhadap aplikasi WhatsApp dengan menggunakan algoritma *K-Means*.
2. Penelitian ini dapat memberikan wawasan yang lebih baik tentang tingkat kepuasan pengguna terhadap aplikasi WhatsApp di Indonesia, yang dapat digunakan oleh perusahaan untuk meningkatkan pengalaman pengguna.
3. Menyediakan pemahaman mendalam tentang jumlah *cluster* yang optimal untuk mengetahui kepuasan pengguna, membantu perusahaan dalam merancang strategi untuk meningkatkan kualitas layanan.
4. Peneliti dapat mengembangkan keterampilan analisis data dengan menerapkan algoritma *K-Means* pada data kepuasan pengguna.
5. Peneliti dapat memperoleh pengalaman penelitian yang berharga, terutama dalam pengumpulan, analisis, dan interpretasi data.
6. Meningkatkan pengetahuan bagi peneliti serta untuk penelitian berikutnya.