

BAB I

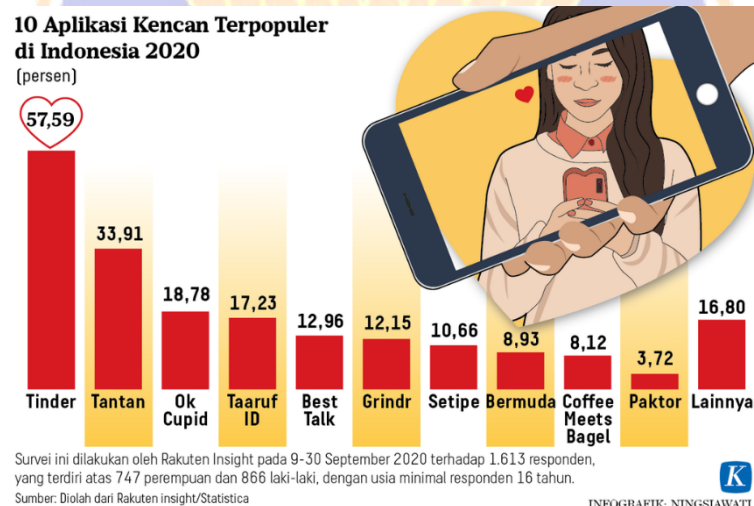
PENDAHULUAN

1.1 Latar Belakang Masalah

Di era yang serba digital masyarakat semakin bebas dalam menggunakan aplikasi. Semakin banyak jenis aplikasi yang tersedia untuk mencari transportasi, belanja, hingga untuk mencari pasangan. Banyak orang yang menggunakan aplikasi kencan *online* atau yang biasa disebut *dating app* dalam mencari pasangan (Waluyo & Revianti, 2019). Pencarian pasangan telah mengalami perubahan seiring dengan hadirnya aplikasi kencan online. Jika dulu orang harus berinteraksi secara langsung untuk mengenal orang baru, namun kini pola perjodohan semakin mengarah ke arah digital dengan memanfaatkan teknologi. Hal ini memungkinkan interaksi yang lebih efisien dan cepat melalui fitur chatting yang tersedia di aplikasi kencan online (Dwi Safitri dkk., 2022). Berdasarkan survey yang dilakukan oleh kumparan.com, dari 619 responden periode 17-31 Agustus 2022, 42,33% pengguna menggunakan *dating app* untuk mencari *friend with benefit* (FWB), 23,26% untuk mencari teman, 19,39% untuk mencari suami/istri, dan 15,02% untuk mencari pacar (Adrian, 2022). Berdasarkan artikel pada antaranews.com (Andarningtyas, 2022), dalam survey “*Mapping a secure path for the future of digital payments in APAC*” yang dilakukan oleh Kaspersky menemukan bahwa hampir seper 2 orang (45%) orang di Asia Tenggara kehilangan uang karena penipuan dari kencan *online*. Penduduk yang pernah menjadi korban penipuan kurang dari 100 USD berjumlah 22% adalah rata-rata Asia Tenggara. Berdasarkan data vertikal Kepolisian Republik Indonesia jumlah kasus penipuan yang terjadi di Indonesia mengalami peningkatan dari tahun 2020 sebesar 561 kasus sampai tahun 2022 sebesar 789 kasus penipuan.

Berdasarkan artikel pada dataindonesia.id, masyarakat Indonesia rela untuk menghabiskan biaya yang lumayan besar untuk berlangganan *dating app*, total pengeluaran konsumen Indonesia untuk aplikasi kencan online adalah \$23,66 juta pada

tahun 2022. Jumlah tersebut meningkat 45,15% dibandingkan tahun lalu, yaitu \$16,3 juta (Rizaty, 2023). Selain itu berdasarkan infografik tahun 2020 (Gambar 1.1) yang dibahas oleh Kompas.id (Wahyudi, 2022), didapatkannya 1613 responden menggunakan *dating app*. Pengguna terbanyak adalah kalangan laki - laki sebanyak 866 responden. Aplikasi yang paling sering digunakan oleh kalangan masyarakat khususnya Indonesia adalah Tinder, dari 57 juta pengunduh aplikasi tinder di dunia, 3,1 juta adalah pengunduh dari Indonesia. Berdasarkan survey tersebut dapat disimpulkan meski terdapat banyak kasus yang muncul akibat *dating app*, masyarakat Indonesia cukup berminat menggunakan *dating app* tersebut. Berdasarkan hal tersebut perlu dilakukan analisis sentiment untuk mengetahui tanggapan masyarakat Indonesia terhadap penggunaan aplikasi *dating online* dari berbagai opini. Menganalisis sentimen sebenarnya dapat dilakukan secara manual, namun karena jumlah data yang sangat besar, maka semakin lama dan usaha yang dilakukan untuk mengklasifikasikan opini tersebut, sehingga perlu diterapkan metode *machine learning* dalam mengklasifikasi opini dari data yang banyak. Salah satu *social network* yang dapat dimanfaatkan untuk melakukan analisis sentiment adalah *Twitter*. *Twitter* memiliki salah satu fitur yakni masyarakat dapat bebas untuk mengekspresikan opini mereka dalam bentuk teks yang disebut *tweets* (Nurfauzan & Maharani, 2021).



Gambar 1.1 Infografik Dating App Terpopuler di Indonesia tahun 2020

Metode yang paling umum digunakan pada saat ini dalam melakukan sentimen adalah metode C4.5, *Support Vector Machine* (SVM) dan *Naïve Bayes*. Algoritma yang cocok untuk klasifikasi sentimen dari ketiga algoritma tersebut adalah Igoritma *Naive Bayes*. Hal ini didukung dengan penelitian dari (Sholikha dkk., 2020) dan (Fikri dkk., 2020) yang membandingkan metode *Naïve Bayes* dengan C4.5 dan SVM yang menghasilkan bahwa metode *Naïve Bayes* lebih unggul untuk melakukan klasifikasi sentimen. Hal ini dapat menjadi alasan *Naïve Bayes* paling cocok dalam melakukan klasifikasi sentimen dikarenakan metode ini mudah dan cepat untuk dipahami, terlebih dalam hal perhitungan hanya memerlukan sedikit data training (Rahman dkk., 2020). Metode *Naïve Bayes* tentunya terdapat kekurangan yakni sangat sensitif dalam memilih fitur. Semakin Banyaknya fitur tidak hanya dapat meningkatkan waktu perhitungan tetapi juga dapat menurunkan *performance* klasifikasi (Negara dkk., 2020). Secara umum, tingkatan lain yang ditemukan dalam klasifikasi untuk membantu meningkatkan akurasi dan menjadi lebih efektif serta efisien yakni dengan melakukan seleksi fitur.

Proses meningkatkan akurasi dari metode *Naïve Bayes* dapat menggunakan *feature selection Information Gain*, tujuan penggunaan seleksi fitur *Information Gain* adalah meningkatkan akurasi klasifikasi dengan mengurangi ukuran dimensi pada data. Menurut (Negara dkk., 2020) metode *Information Gain* dapat mengoptimasi nilai akurasi serta mengurangi jumlah fitur analisis sentimen. Penelitian yang dilakukan oleh Negara dkk., analisis sentiment dengan mengkombinasikan metode *Naïve Bayes Classifier* dan *Information Gain*. Hasil dari penelitian tersebut menunjukkan perbandingan antara penggunaan metode *Information Gain*, dengan akurasi lebih tinggi dangan menggunakan *Information Gain* daripada tanpa menggunakan *Information Gain*. Berdasarkan permasalahan yang sudah diuraikan, maka penelitian ini akan melakukan perbandingan analisis sentimen masyarakat terhadap penggunaan aplikasi *dating online* setelah mengetahui kasus yang terjadi dan diiringi peningkatan penggunaan aplikasi kencan online menggunakan metode *Naive Bayes Classifier* disertai *Information Gain* dengan judul “**Analisis Sentimen Penggunaan Aplikasi**

Dating Online Melalui Twitter Menggunakan Metode Naïve Bayes Classifier Dan Information Gain

1.2 Rumusan Masalah

Berdasarkan permasalahan yang sudah diuraikan sebelumnya, maka dapat dirumuskan masalah seperti berikut:

1. Bagaimana implementasi kombinasi *Naive Bayes Classifier* dengan *Information Gain* untuk menganalisis sentimen komentar terhadap aplikasi *dating online*?
2. Bagaimana perbandingan tingkat performa algoritma *Naive Bayes Classifier* dengan seleksi fitur *Information Gain* dan tanpa menggunakan *Information Gain* untuk menganalisis sentimen komentar terhadap aplikasi *dating online*?

1.3 Tujuan Penelitian

1. Untuk mengetahui bagaimana implementasi kombinasi algoritma *Naive Bayes Classifier* dan seleksi fitur *Information Gain* dalam menganalisis sentimen masyarakat terhadap aplikasi *dating online*
2. Untuk mengetahui perbandingan tingkat performa algoritma *Naive Bayes Classifier* dengan seleksi fitur *Information Gain* dan tanpa menggunakan *Information Gain* untuk menganalisis sentimen komentar terhadap aplikasi *dating online*

1.4 Batasan Masalah

1. Dataset pada penelitian ini merupakan data komentar berbahasa Indonesia yang diperoleh dari media *twitter*.
2. Dataset yang digunakan akan menggunakan keyword "*dating online*", "*kencan online*", "*datingapp*", "*tinder*", "*tantan*", "*okcupid*" yang dikumpulkan dari periode Desember tahun 2020 sampai Januari tahun 2022.
3. Pada penelitian ini, analisis sentimen yang dilakukan berfokus pada penggunaan aplikasi *dating online* secara umum.
4. Pada penelitian ini tidak dipertimbangkan masalah *imbalanced data*

1.5 Manfaat Penelitian

Manfaat pada penelitian ini adalah sebagai berikut:

1. Manfaat Penulis

Dapat mengimplementasikan dan mengetahui performa algoritma *Naive Bayes Classifier*.

2. Manfaat Pembaca

Mengetahui sudut pandang masyarakat mengenai dampak dan manfaat mengenai penggunaan aplikasi *dating online* serta dapat digunakan untuk pengambilan keputusan apakah aplikasi *dating online* ini berdampak positif dalam kehidupan masyarakat.

3. Manfaat Akademik

Sebagai referensi dalam analisis sentimen dan pertimbangan untuk menindak lanjuti penelitian yang relevan sekaligus sebagai referensi di perpustakaan Universitas Pendidikan Ganesha bagi penelitian relevan di masa yang akan datang.

