

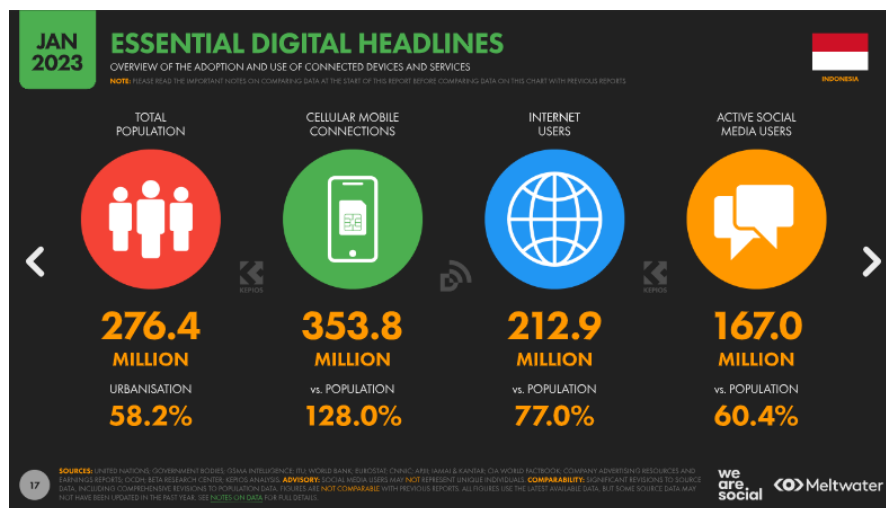
BAB I

PENDAHULUAN

1.1 Latar Belakang

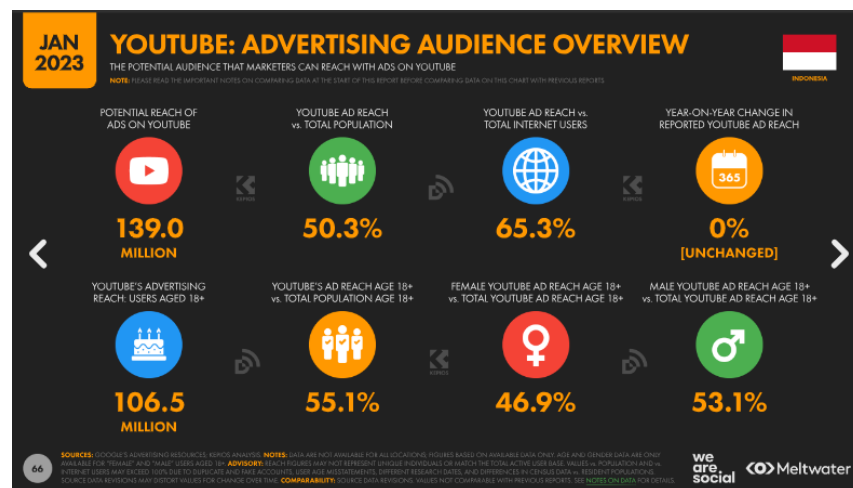
Perkembangan teknologi masa kini, kemajuan dalam teknologi informasi dan komunikasi berlangsung dengan cepat, dan ini terlihat jelas melalui kemudahan akses internet. Salah satu hasil terkemuka dari perkembangan dalam teknologi informasi dan komunikasi adalah pemanfaatan internet sebagai sarana utama dalam berkomunikasi (Wirany, Natasha, and Kurniawan 2022).

Berdasarkan (<https://datareportal.com/reports/digital-2023-indonesia> diakses pada 23 Oktober 2023) survei data We Are Social : Data pengguna internet Indonesia pada bulan Januari 2023 mencapai 212.900.000 jiwa dari 77.0% total populasi. Sementara pengguna media sosial sebesar 167.000.000 jiwa dengan populasi 60.4%. Total jumlah penduduk Indonesia pada bulan Januari 2023 adalah sebesar 276.400.000 jiwa. Waktu yang dihabiskan untuk mengakses di internet rata-rata per hari adalah selama 7 jam, 42 menit penggunaan internet.



Gambar 1. 1 Data Pengguna Internet dan media sosial Januari 2023 di Indonesia

Salah satu media sosial yang digemari oleh anak muda dalam mencari berbagai jenis konten video adalah *platforms* Youtube (Salim and Sutabri 2023). Menurut (<https://datareportal.com/reports/digital-2023-indonesia> html diakses pada 23 Oktober 2023) jumlah pengguna *platforms* Youtube pada Januari menurut laporan data *We Are Social* telah mencapai sebesar 139.000.000 jiwa, yang sudah menggunakan *platforms* Youtube pada bulan Januari 2023. Dengan total jumlah penduduk Indonesia pada bulan Januari 2023 sebanyak 276.400.000 jiwa.



Gambar 1. 2 Data Penggunaan *Platform* Youtube Januari 2023 di Indonesia

Dibalik pertumbuhan penggunaan internet yang pesat, terdapat fakta yang mengkhawatirkan, yaitu peningkatan yang signifikan dalam kasus *cyberbullying*. Menurut catatan KPAI, dalam kurun waktu sembilan tahun, mulai dari 2011 hingga 2019, terdapat 37.381 pengaduan terkait kekerasan terhadap anak. Kasus *bullying*, baik yang terjadi di lingkungan pendidikan maupun di media sosial, mencapai angka 2.473 laporan, dan trennya terus meningkat (<https://www.kpai.go.id/publikasi/sejumlah-kasus-bullying-sudah-warnaicatat-masalah-anak-di-awal-2020-begini-kata-komisioner-kpai> html diakses pada 23 Oktober 2023).

Cyberbullying atau perundungan online, merupakan fenomena yang semakin meningkat seiring dengan perkembangan teknologi informasi dan media sosial. Dalam kasus ini, individu menggunakan berbagai platform digital seperti aplikasi media sosial, pesan instan, dan forum online untuk melakukan intimidasi, pelecehan, atau ancaman kepada orang lain. Tindakan tersebut dimaksudkan untuk mempermalukan, mengintimidasi, menyebar keburukan dan kebencian di media siber atau dunia maya, baik ditujukan secara khusus kepada korban maupun dengan cara diketahui publik (Etd n.d.). Dalam beberapa kasus, *cyberbullying* dapat berdampak serius pada kesejahteraan mental dan emosional korban, bahkan dapat menyebabkan dampak jangka panjang yang merugikan.

Penelitian terdahulu yang menjadi acuan dalam penelitian ini adalah penelitian yang dilakukan oleh (Maulana et al. 2020) yang berjudul “Analisa sentiment *cyberbullying* di jejaring sosial twitter dengan algoritma *naïve bayes*”. Hasil pengujian dengan data uji *real time* pada tanggal 12 Mei 2020 pukul 01.00 WIB mendapatkan nilai akurasi sebesar 76%. Metode ini cukup baik dalam mengklasifikasikan tweet positif dan negatif, namun pada proses pengujian penelitian ini dalam mendeteksi tweet yang mengandung unsur *cyber bullying* masih kurang baik, dikarenakan masih terdapatnya tweet yang tidak mengandung unsur *cyber bullying* didalam data latih yang memiliki label tweet negatif.

Adapula penelitian yang dilakukan oleh (Purnamasari et al. 2018) yang berjudul “Identifikasi tweet *cyberbullying* pada aplikasi twitter menggunakan metode *support Vector Machine (SVM)* dan *information gain (IG)* sebagai seleksi fitur”. Hasil yang didapatkan dengan metode SVM adalah *accuracy* 75%, *precision* 70,27%, *recall* 86,66% dan *f-measure* 77,61% pada percobaan nilai $\text{iterMax} = 20$, $\lambda = 0,5$, $\gamma = 0,001$, $\gamma = 0,000001$, dan $C = 1$. Nilai threshold

terbaik seleksi fitur *information gain* adalah 90%, dengan nilai *accuracy* 76,66%, *precision* 72,22%, *recall* 86,66% dan *f-measure* 78,78%.

Kemudian dilihat dari penelitian terdahulu Penelitian ini akan berfokus pada *platform* YouTube, sebuah *platform* digital yang populer di kalangan pengguna internet. Dalam beberapa bulan terakhir, akun YouTube *Food Vlogger* Mgdalenaf mengalami berbagai komentar negatif dari pengguna YouTube lainnya. Dalam wawancara di saluran YouTube milik Samuel Christ, Magdalena secara terbuka menceritakan pengalaman kurang menyenangkan saat menjadi *Food Vlogger* Indonesia. Dia mengungkapkan bahwa dalam suatu kejadian, pedagang tidak menyadari seberapa besar pengaruh dan jumlah pengikutnya selama ini. Hal ini membuat publik menafsirkan bahwa Magdalena seolah-olah menginginkan perlakuan istimewa dan enggan membayar makanan yang dibelinya karena merasa terkenal.

Selanjutnya penelitian ini menggunakan Metode *Naïve Bayes*, dengan teknik pengambilan data komentar melalui proses *crawling* menggunakan Google Colab *Research Python*. Diharapkan hasil penelitian ini dapat memberikan pemahaman kepada pengguna *platform* YouTube mengenai apakah suatu komentar di *platform* tersebut mengandung unsur *cyberbullying* atau tidak. Selain itu, diharapkan juga memberikan kontribusi penting dalam upaya mengatasi kasus-kasus *cyberbullying* di *platform* digital secara lebih efektif. Maka, penelitian ini berjudul **“DETEKSI CYBERBULLYING PADA KOMENTAR YOUTUBE DENGAN METODE NAÏVE BAYES (Studi Pada Akun Youtube *Food Vlogger* Mgdalenaf)”**, yang menjadi fokus utama dari penelitian ini."

1.2 Ruang Lingkup

1.1.1 Ruang Lingkup Subjek

Ruang lingkup subjek dalam penelitian ini adalah pada *platform* Youtube ini akan fokus pada identifikasi teks yang mencurigakan terkait *cyberbullying* di dalam akun Youtube *Food Vlogger Mgdalenaf*.

1.1.2 Ruang Lingkup Objek

Ruang lingkup objek dalam penelitian ini melibatkan pengumpulan data dan pemrosesan data teks yang mencakup komentar atau konten yang mungkin mengandung perilaku *cyberbullying* atau tidak.

1.1.3 Ruang Lingkup Waktu

Waktu yang digunakan dalam melaksanakan penelitian ini adalah Oktober 2023 sampai dengan Februari 2024.

1.1.4 Ruang Lingkup Ilmu Pengetahuan

Ruang lingkup ilmu penelitian ini mencakup pemahaman tentang *Text Mining*, dengan pemanfaatan metode *Naïve Bayes*. Selain itu, penelitian ini juga mencakup identifikasi *cyberbullying*, implementasi dalam *platform* Youtube, dan aspek etika serta privasi dalam penanganan data teks.

1.3 Rumus Masalah

Berdasarkan latar belakang di atas, perumusan masalah ini bertujuan untuk upaya membatasi penelitian agar lebih terarah dan tidak terlalu luas namun tetap dalam fokus yang diharapkan dan yang telah di tentukan, maka rumusan masalah yang akan peneliti angkat sebagai berikut :

- a. Bagaimana karakteristik *cyberbullying* dalam komentar video YouTube pada akun *Food Vlogger Mgdalenaf*?

- b. Apakah Metode *Naïve Bayes* efektif dalam mendeteksi dan mengklasifikasikan komentar-komentar yang bersifat *cyberbullying*?

1.4 Batasan Masalah

- a. Penelitian ini terbatas pada *platform* YouTube sebagai lingkungan utama untuk analisis *cyberbullying*.
- b. Analisis *cyberbullying* terbatas pada komentar-komentar yang ditemukan di akun YouTube *Food Vlogger* Mgdalenaf.
- c. Dataset komentar untuk analisis diperoleh melalui proses *crawling* menggunakan Google Colab *Research Python*.
- d. Dataset yang digunakan dari tanggal 05 April – 11 November 2023

1.5 Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah di atas, terdapat beberapa tujuan dalam penelitian ini adalah sebagai berikut:

- a. Menganalisis karakteristik *cyberbullying* dalam komentar video YouTube.
- b. Mengukur efektivitas metode *Naïve Bayes* dalam mengidentifikasi dan mengklasifikasikan komentar yang bersifat *cyberbullying*.

1.6 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah:

1.5.1 Bagi Penulis

Dari penelitian ini, penulis memiliki kesempatan untuk mengembangkan ilmu pengetahuan dalam bidang *Text Mining* yang meliputi pengelolaan data *text* dari *platform* Youtube dengan memanfaatkan metode *Naïve Bayes* pada penggunaan *open-source* Google Colab *Phyton*.

1.5.2 Bagi Institusi

Hasil penelitian ini dapat digunakan sebagai bahan acuan bagi peneliti berikutnya yang berminat dalam mengembangkan strategi pengenalan dan penanganan *cyberbullying* dalam konteks *platform* Youtube. Dengan demikian, penelitian ini dapat memberikan kontribusi penting untuk memahami dan mengatasi isu-isu keamanan *cyber* dan perilaku negatif dalam lingkungan digital, serta memberikan landasan bagi penelitian lebih lanjut dalam menghadapi tantangan ini.

1.5.3 Bagi Pengguna YouTube

Penelitian ini memberikan manfaat bagi pengguna YouTube dengan memberikan dampak positif. Melalui deteksi *cyberbullying* di platform YouTube menggunakan Metode *Naïve Bayes*, penelitian ini bertujuan untuk mengetahui apakah suatu komentar termasuk dalam kategori *cyberbullying* atau *non-cyberbullying*.

1.7 Sistematika Penulisan

BAB I PENDAHULUAN

Dalam pendahuluan tercantum antara lain latar belakang, rumusan masalah, ruang lingkup, tujuan penelitian, manfaat penelitian dan sistematika penulisan tentang *cyberbullying text mining* pada *platform* Youtube dengan metode *naïve bayes*.

BAB II LANDASAN TEORI

Dalam bab ini berisi tentang teori yang didapat dari berbagai sumber sebagai sebuah landasan atau acuan untuk klasifikasi kasus-kasus *cyberbullying* pada *platform* Youtube, dengan mengimplementasikan metode *Naïve Bayes* dan melakukan klasifikasi data lebih mudah menggunakan Google Colab *Phyton*.

BAB III METODE PENELITIAN

Dalam bab ini berisi metode pengumpulan data dengan *crawling*, implementasikan metode *Naïve Bayes* klasifikasi data menggunakan Google Colab *Phyton* dan pendekatan penyelesaian masalah yang dinyatakan dalam perumusan masalah pada penelitian yang dilakukan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas hasil klasifikasi kasus *cyberbullying* pada *platform* YouTube yang telah dianalisis dan diimplementasikan menggunakan metode *Naïve Bayes*.

BAB V SIMPULAN DAN SARAN

Dalam bab ini berisi tentang kesimpulan dari keseluruhan penulisan skripsi ini dan juga saran yang dapat membantu perbaikan dan pengembangan selanjutnya.