

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Pada tahun 2022, basis pengguna *YouTube* mencapai 2,56 miliar pengguna dari seluruh dunia dan angka ini diproyeksikan akan tumbuh mencapai 2,85 miliar pengguna pada tahun 2025 (Ceci, YouTube - Statistics & Facts, 2022). Semenjak diluncurkan pada tahun 2005 dan diakuisisi oleh *Google*, *YouTube* menjadi platform video online terbesar di dunia dengan menampilkan beragam konten media yang mencakup konten video korporat, video musik, video promosi produk, video blog, video edukasi. Meningkatnya pengguna berbanding lurus dengan bertambah banyaknya konten video yang diunggah di *YouTube*. Di Indonesia sendiri pengguna *YouTube* mencapai 139 juta pada bulan Januari 2022 (Kemp, 2022).

Beragam reaksi opini dalam komentar pengguna ketika melihat video yang diunggah di *YouTube* sangat memengaruhi reputasi video tersebut terlepas dari bagaimana opini masyarakat dalam mengemukakan opini tersebut. Dalam menilai konten video terdapat dua ulasan yaitu ulasan berdasarkan jumlah *like* dan ulasan berupa komentar tertulis yang disematkan pada konten video tersebut. Ulasan berupa jumlah *like* menunjukkan evaluasi pengguna dalam melihat konten video dalam angka, namun ulasan berupa komentar tertulis mampu bercerita lebih mendalam tentang pengalaman pengguna ketika melihat konten video yang diunggah oleh pembuat konten. Dengan banyaknya komentar membuat pembuat konten mengalami kesulitan

dalam mencari tahu dan memutuskan apakah konten video diterima dengan baik oleh pengguna *YouTube*.

Konferensi Tingkat Tinggi G20 2022 Bali (KTT G20 2022) merupakan sebuah forum utama kerja sama ekonomi internasional yang beranggotakan 19 negara dengan ekonomi terbesar di dunia dan 1 negara Uni-Eropa. KTT G20 2022 juga merupakan forum untuk membangun kemitraan dan kerja sama untuk mengatasi masalah-masalah global (Samad & Permatasari, 2023). KTT G20 membahas isu-isu seperti Hubungan Rusia dengan Ukraina yang menimbulkan topik Presiden Rusia akan dikeluarkan dari keanggotaan G20, penanganan pandemi covid-19 yang melanda perekonomian global (Indah, 2022).

Memahami sesuatu yang dipikirkan oleh orang lain akan menjadi informasi yang sangat penting bagi pembuat konten untuk membuat video yang dapat diterima baik oleh pengguna *YouTube*. Sehingga dapat dibuat sebuah pendekatan untuk mengetahui bagaimana persepsi pengguna *YouTube* terhadap suatu video dengan menggunakan data analisis sentimen yang didapat dari komentar tekstual. Pendekatan *text mining* dengan *Nature Language Processing* adalah salah satu alternatif terbaik dalam memberikan makna dari setiap komentar (Muhammad, Bukhori, & Pandunata, 2019). Pengklasifikasian sentimen dibagi menjadi dua kategori yaitu positif dan negatif dan bertujuan untuk melihat opini, sikap, emosi, dan serta penilaian masyarakat terhadap suatu produk, individu, peristiwa, dan topik (Afdhal, et al., 2022). Dengan adanya pendekatan ini, banyak manfaat yang diberikan dibandingkan dengan hanya menghitung jumlah *like* dan membaca satu per satu komentar yang ada pada channel youtube, dan dapat secara mudah mendapatkan hasil penilaian dari video yang telah

diunggah. Pengurangan tenaga dan waktu yang dikeluarkan karena data dapat diambil secara cepat dan mudah melalui teknik *crawling* di *YouTube* dengan menggunakan *YouTube API* (Application Programming Interface) untuk mendapatkan data yang bersangkutan berupa komentar, jumlah *like* dan jumlah *views* pada sebuah video (Baravkar, Jaiswal, & Chhoriya, 2020).

Penelitian ini melakukan analisis sentimen pengguna *YouTube* dalam keberlangsungan G20 yang diadakan pada tanggal 15-16 November dengan memberikan pandangan dari beberapa *channel youtube* yang dipilih menggunakan metode *Naïve Bayes Classifier* dan *Support Vector Machine*. Kedua metode ini digunakan karena memiliki kesamaan yaitu pada jenis model *Linear Classifier*. Metode *Naïve Bayes Classifier* (NBC) merupakan algoritma yang dapat mengklasifikasi suatu variabel tertentu dengan menggunakan metode probabilitas dan statistik (Khotimah & Utami, 2022). Metode *Support Vector Machine* (SVM) adalah metode klasifikasi dengan prinsip mencari *hyperplane* yang memiliki margin terbesar. *Hyperplane* memisahkan data antar kelas/kategori sedangkan *margin* merupakan jarak antara *hyperplane* dengan data terdekat (Irmada & Astriratma, 2021).

Penelitian ini dilakukan dengan mengambil data komentar dari video youtube pada acara siaran langsung yang diadakan pada KTT G20 2022 Bali pada tanggal 15-16 November 2022. Hasil penelitian ini yaitu hasil analisis sentimen, perbandingan antara metode *Naïve Bayes* dan *Support Vector Machine*, dan juga *website* yang dapat melakukan prediksi sentimen terhadap komentar *youtube*. Dari hasil tersebut diharapkan dapat mengetahui sentimen masyarakat terhadap KTT G20 dan juga dapat

memprediksi sentimen sehingga dapat membantu pihak-pihak yang ingin mengetahui kata kunci yang sama.

Berdasarkan permasalahan yang ada maka penelitian yaitu Analisis Sentimen Komentar Video Youtube Konferensi Tingkat Tinggi G20 2022 Menggunakan Metode Naïve Bayes dan Support Vector Machine.

1.2 Rumusan Masalah

Berdasarkan uraian yang disampaikan dalam latar belakang. Adapun rumusan masalah dalam penelitian ini adalah:

1. Bagaimana melakukan analisis sentimen dalam mengklasifikasikan komentar video *YouTube* dengan metode *Naïve Bayes* dan *Support Vector Machine*?
2. Bagaimana mengetahui metode yang lebih optimal antara *Naïve Bayes* dan *Support Vector Machine* dalam melakukan klasifikasi komentar *YouTube*?
3. Bagaimana melakukan prediksi sentimen terhadap komentar *youtube* berbasis *website*?
4. Bagaimana memvisualisasikan hasil dari klasifikasi sentimen berbasis *website*?

1.3 Batasan Masalah

Agar tidak terjadi penyimpangan dalam proses penelitian dan pengembangan sistem, maka ditetapkan batasan masalah dalam penelitian ini. adapun batasan masalah diantaranya sebagai berikut:

1. Video yang diambil yaitu video KTT G20 2022 pada rentang waktu 15-16 November 2022 berdasarkan *rating* video yang paling banyak ditonton.

2. Data yang digunakan diambil menggunakan *YouTube API (Application Programming Interface)* dengan *API key* penulis.
3. Data komentar dalam penelitian ini yang diambil menggunakan Bahasa Indonesia yang benar.
4. Komentar yang didapat diklasifikasikan ke dalam dua analisis sentimen yaitu positif dan negatif secara otomatis dengan VADER.
5. Komentar yang diberi label tidak dapat membedakan termasuk sarkasme atau tidak.
6. Algoritma yang digunakan dalam pengklasifikasian adalah *Naïve Bayes Classifier* dan *Support Vector Machine*.
7. Bahasa pemrograman yang digunakan adalah Bahasa Pemrograman Python.

1.4 Tujuan Penelitian

Tujuan dalam penelitian ini merupakan jawaban dari rumusan masalah yang telah disampaikan. Adapun tujuan penelitian ini adalah:

1. Mengetahui tanggapan pengguna *YouTube* bersifat positif atau negatif yang berasal dari video youtube KTT G20 2022.
2. Melakukan klasifikasi dari metode *Naïve Bayes Classifier* dan *Support Vector Machine* dan mengetahui dari kedua metode yang lebih efektif dalam melakukan analisis sentimen.
3. Membuat suatu aplikasi yang dapat melakukan prediksi sentimen terhadap komentar video youtube berbasis website.

4. Membuat visualisasi dari klasifikasi data dan menampilkan kata yang menjadi pemicu sentimen positif dan negatif dalam beropini dan mengimplementasikan ke dalam suatu aplikasi berbasis website.

1.5 Sistematika Penulisan

Sistematika penulisan skripsi ini akan membantu mengarahkan penyusunan laporan supaya tidak menyimpang dan sebagai acuan dalam mencapai tujuan penulisan laporan skripsi sesuai dengan apa yang diharapkan. Langkah-langkah dalam proses penyusunan skripsi sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi tentang gambaran umum isi penelitian diantaranya latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini berisi tentang penelitian terdahulu untuk membandingkan penelitian terdahulu dengan penelitian saat ini, dasar teori yang berhubungan dengan masalah yang akan dibahas seperti *text mining*, *youtube*, *pre-processing text*, analisis sentimen, dan klasifikasi teks.

BAB III METODE PENELITIAN

Bab ini berisi tentang metode penelitian yang akan dibuat meliputi studi literatur, analisis kebutuhan, dan perancangan model dan visualisasi.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi tentang hasil dan pembahasan berupa penjelasan penelitian mengenai pengumpulan data, pembangunan model, dan pengujian sistem serta akan membahas hasil pengujian.

BAB V PENUTUP

Bab ini berisi tentang kesimpulan mengenai seluruh hasil penelitian dan berisikan saran terhadap penelitian selanjutnya.

DAFTAR PUSTAKA

Pada bagian ini berisi tentang literatur yang digunakan sebagai pedoman yang membantu pengerjaan skripsi.

LAMPIRAN

Pada bagian ini berisi tentang data atau pelengkap yang menunjang dalam pembuatan skripsi.