

BAB I PENDAHULUAN

1.1 Latar Belakang Masalah

YouTube telah berkembang menjadi platform video terkemuka di dunia, dengan lebih dari 2,5 miliar pengguna aktif setiap bulan [1]. Di Indonesia sendiri, sekitar 94% pengguna internet memanfaatkan YouTube sebagai sumber hiburan, informasi, dan edukasi [2]. Popularitas ini menjadikan YouTube pusat distribusi konten multimedia yang sangat beragam, mulai dari video tutorial, ulasan produk, hingga siaran langsung acara global.

Di tengah pertumbuhan pesat tersebut, muncul berbagai tantangan dalam menilai dan memilih konten secara cepat. Pertama, kebijakan terbaru YouTube yang menghilangkan tampilan jumlah *dislike* mengurangi kemampuan penonton untuk menilai respons audiens secara cepat. Tanpa metrik tersebut, pengguna kesulitan mendapatkan gambaran awal mengenai reputasi dan kualitas suatu video.

Kedua, skala youtube yang sangat besar membuat volume unggahan video yang sangat besar pula. Hal ini memperparah kesulitan dalam memilah video bermutu. Ketiga, menurut [3] banyak video dihasilkan secara massal oleh *content farm* atau dibuat menggunakan kecerdasan buatan (AI) untuk keuntungan finansial semata, tanpa memperhatikan akurasi fakta atau nilai edukatif. Video semacam ini sering kali menyuguhkan informasi dangkal dan mengaburkan konten yang benar-benar bermanfaat. Akibatnya, proses seleksi konten menjadi memakan waktu karena pengguna terpaksa menonton sebagian video atau membaca ratusan komentar secara manual untuk mengevaluasi kualitasnya. Metode konvensional ini tidak hanya tidak efisien, tetapi juga rawan bias.

Untuk menjawab tantangan tersebut, diperlukan sistem yang mampu menganalisis sentimen komentar secara otomatis dan menyajikan ringkasan respons audiens hanya dengan memasukkan tautan video atau kata kunci pencarian. Dengan dukungan fitur transkripsi dan peringkasan video, pengguna dapat memahami inti konten tanpa menonton keseluruhan video. Pendekatan ini diharapkan mempercepat, memudahkan, dan memperjelas proses seleksi konten, sehingga setiap pengguna dapat memperoleh pengalaman menonton yang lebih terarah dan bermakna.

1.2 Rumusan Masalah

Seiring dengan meningkatnya jumlah video yang diunggah ke YouTube setiap hari, pengguna menghadapi tantangan dalam menilai kualitas suatu konten. Salah satu faktor yang memperumit seleksi video adalah penghapusan tampilan jumlah *dislike*, sehingga penonton kesulitan untuk menilai reaksi audiens terhadap suatu video secara cepat. Selain itu, banyaknya video yang diproduksi oleh *content farm* atau berbasis kecerdasan buatan (AI) sering kali hanya mengejar keuntungan finansial tanpa memperhatikan nilai edukatif atau informatif. Pengguna harus membaca komentar secara manual untuk memahami sentimen audiens terhadap suatu video, yang tentu saja tidak efisien mengingat volume data yang sangat besar. Oleh karena itu, diperlukan suatu sistem yang dapat menganalisis sentimen komentar video YouTube secara otomatis guna membantu pengguna dalam menilai kualitas suatu konten. Berdasarkan latar belakang, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana merancang sistem analisis sentimen komentar video YouTube yang dapat menyajikan ringkasan sentimen tanpa perlu membaca seluruh komentar secara manual?
2. Bagaimana penerapan metode atau model yang efektif dalam mengklasifikasikan sentimen komentar video YouTube?
3. Bagaimana kinerja masing-masing metode atau model yang cocok untuk digunakan dalam sistem analisis sentimen secara real-time?

1.3 Batasan Masalah

Dalam proses pengembangan sistem analisis sentimen komentar YouTube, penting untuk menetapkan batasan yang jelas agar ruang lingkup penelitian tetap terarah dan tidak melebar ke aspek yang belum relevan untuk saat ini. Penelitian ini berfokus pada pembuatan sistem yang sederhana namun tetap fungsional, dengan mempertimbangkan keterbatasan perangkat keras, bahasa, serta cakupan pengguna. Sistem ini dirancang bukan untuk menggantikan seluruh proses kurasi konten, melainkan untuk membantu pengguna dalam memilah video berdasarkan tanggapan audiens yang tercermin dari komentar. Menurut [4] istilah "konten berkualitas" tidak didefinisikan secara mutlak, tetapi lebih mengarah pada video yang dianggap lebih layak untuk ditonton

karena memuat tanggapan positif, bersifat informatif, atau memberikan nilai tambah, baik berupa hiburan yang sehat maupun informasi yang bersifat edukatif. Maka dari itu, berikut adalah batasan-batasan yang diterapkan dalam penelitian ini:

1. Ukuran model yang digunakan dibawah IGB.
2. Pengumpulan data dan metode atau model harus kompatibel dengan *framework Flask*.
3. Data yang dianalisis pada sistem jadi hanya berasal dari komentar video YouTube yang diperoleh melalui YouTube API dengan API key milik penulis.
4. Analisis hanya dilakukan terhadap komentar berbahasa Inggris karena keterbatasan model dan pustaka yang digunakan.
5. Sistem tidak memfilter komentar bot, sehingga komentar otomatis masih mungkin ikut dianalisis.
6. Uji coba sistem hanya melibatkan responden dari kalangan terbatas, seperti teman dan lingkungan sekitar peneliti.
7. Sentimen dibagi menjadi tiga kategori, yaitu Positive, Neutral, dan Negative, untuk menjaga kesederhanaan sistem.

Dengan adanya batasan ini, sistem yang dikembangkan diharapkan tetap realistis untuk dijalankan dalam lingkungan terbatas dan mampu memberikan hasil yang cukup representatif dalam membantu pengguna menilai konten video YouTube secara lebih cepat dan informatif.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk membuat sebuah sistem berbasis web yang bisa menganalisis sentimen dari komentar-komentar video YouTube secara otomatis. Ide dasarnya cukup sederhana: supaya orang tidak perlu membaca komentar satu per satu hanya untuk mengetahui video itu disukai atau tidak. Dengan sistem ini, pengguna cukup memasukkan link video atau mencari berdasarkan kata kunci, lalu bisa langsung melihat ringkasan sentimen audiens secara cepat.

Secara khusus, sistem ini ditujukan untuk memberikan kemudahan dalam proses seleksi konten yang dianggap lebih relevan dan bermanfaat bagi pengguna. Meskipun tidak secara eksplisit menilai kualitas isi video, sistem ini memanfaatkan opini publik

sebagai indikator awal untuk mengidentifikasi video yang berpotensi memberikan pengalaman menonton yang lebih bernilai. Dalam konteks ini, fokus penelitian tidak hanya berhenti pada aspek teknis klasifikasi komentar, tetapi diharapkan sistem bisa membantu pengguna memilah video yang mungkin lebih bermanfaat atau layak ditonton, terutama yang mengandung unsur edukatif.

Video YouTube yang mengandung unsur edukatif umumnya memperoleh komentar dengan sentimen yang lebih positif. Hal ini dibuktikan dalam penelitian [5], yang menemukan bahwa sebagian besar komentar pada video edukatif—baik di bidang STEM maupun non-STEM—didominasi oleh sentimen positif dan netral, dengan komentar negatif yang jauh lebih sedikit. Dukungan tambahan juga muncul dari penelitian [6], yang menunjukkan bahwa video edukasi dari kanal YouTube seperti TEDx dan NeetCode secara umum memperoleh komentar dengan distribusi sentimen positif dan netral yang tinggi, serta dianggap memberikan pengalaman menonton yang tidak memicu respons emosional negatif. Demikian, penelitian ini memiliki beberapa tujuan utama yang menjawab rumusan masalah pada sub bab 1.2. Tujuan dari penelitian ini yaitu:

1. Mengembangkan sistem analisis sentimen komentar YouTube berbasis web.
2. Menyediakan ringkasan hasil analisis sentimen dalam bentuk yang mudah dipahami sebagai gambaran umum terhadap suatu video.
3. Membantu pengguna dalam menyaring dan mengevaluasi konten video yang dianggap berkualitas, khususnya yang memiliki kecenderungan edukatif, berdasarkan opini penonton lain.
4. Membandingkan dan mengevaluasi performa beberapa metode analisis sentimen seperti untuk melihat mana yang paling sesuai digunakan dalam sistem.
5. Menentukan model terbaik berdasarkan metrik evaluasi seperti *F1-Score*, serta mempertimbangkan efisiensi dalam konteks penggunaan real-time.
6. Mengimplementasikan model tersebut ke dalam sistem yang dapat digunakan secara langsung melalui antarmuka web yang sederhana.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat yang luas bagi berbagai pihak yang terlibat dalam ekosistem YouTube. Sistem analisis sentimen yang dikembangkan tidak hanya berguna bagi pengguna dalam menyeleksi konten berdasarkan sentimen, tetapi juga dapat membantu konten kreator memahami respons audiens

terhadap video yang mereka buat. Selain itu, bagi penulis sendiri, penelitian ini menjadi sarana untuk mengembangkan kemampuan dalam bidang teknologi pemrosesan bahasa alami dan rekayasa perangkat lunak berbasis web. Secara lebih rinci, manfaat penelitian ini dapat dibagi menjadi dua kategori:

a. Manfaat Akademis:

- 1) Memberikan kontribusi dalam pengembangan sentimen analisis pada sistem realtime, khususnya dalam konteks media sosial YouTube.
- 2) Menambah wawasan terkait pemilihan dan evaluasi model analisis sentimen berdasarkan metrik *F1-Score* dan efisiensi waktu.
- 3) Menjadi referensi bagi peneliti atau mahasiswa yang ingin melakukan penelitian lanjutan di bidang pemrosesan bahasa alami dan sistem informasi berbasis prototyping.

b. Manfaat Praktis:

- 1) Membantu pengguna YouTube dalam mengevaluasi konten video secara cepat berdasarkan analisis komentar otomatis.
- 2) Memberikan wawasan bagi konten kreator mengenai persepsi atau reaksi audiens terhadap konten yang mereka unggah.
- 3) Menghadirkan sistem yang dapat menghemat waktu pengguna dalam proses pemilihan konten, tanpa perlu membaca komentar satu per satu.
- 4) Menjadi dasar untuk pengembangan sistem serupa yang lebih luas dan dapat digunakan secara publik.
- 5) Memberikan pengalaman langsung kepada penulis dalam membangun sistem nyata yang dapat diuji secara fungsional.

1.6 Sistematika Penulisan

Skripsi ini disusun secara sistematis dalam lima bab utama yang saling terhubung untuk membentuk alur penelitian yang utuh, mulai dari identifikasi masalah hingga kesimpulan dan saran. Sistematika ini bertujuan untuk mempermudah pembaca dalam memahami konteks penelitian, pendekatan yang digunakan, proses pengembangan sistem, serta hasil evaluasi dari model analisis sentimen yang diterapkan. Setiap bab dirancang untuk menjawab aspek-aspek penting dari penelitian, sekaligus menunjukkan kontribusi sistem yang dikembangkan dalam menjawab permasalahan nyata di lapangan.

Berikut adalah bab-bab tersebut:

BAB I PENDAHULUAN membahas latar belakang permasalahan yang melandasi penelitian, yaitu hilangnya indikator *dislike* dan meningkatnya konten tidak berkualitas di YouTube, yang menyulitkan pengguna dalam menilai kualitas video. Bab ini juga memuat rumusan masalah, tujuan, manfaat, batasan masalah, serta penjelasan sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA, menguraikan teori-teori yang relevan, seperti analisis sentimen, teknik *preprocessing* teks, algoritma klasifikasi (Multinomial *Naïve Bayes* dan *Linear SVC*), serta metode leksikon (VADER dan TextBlob). Selain itu, bab ini juga menyajikan tinjauan penelitian terdahulu sebagai dasar dan perbandingan.

BAB III METODOLOGI PENELITIAN, menjelaskan pendekatan prototyping yang digunakan, dimulai dari pengumpulan komentar melalui YouTube API, *preprocessing* teks, pelabelan data, pelatihan model klasifikasi, hingga evaluasi sistem. Diagram alur kerja dan perancangan antarmuka juga dijelaskan dalam bab ini.

BAB IV HASIL DAN PEMBAHASAN, menyajikan hasil implementasi dan pengujian sistem. Pembahasan mencakup performa model berdasarkan *F1-Score*, hasil evaluasi metode leksikon, serta penilaian terhadap tampilan dan fungsionalitas sistem *frontend* dan *backend* dari dua iterasi prototipe.

BAB V PENUTUP, berisi kesimpulan dari hasil penelitian, termasuk perbandingan performa model, keberhasilan sistem dalam menyajikan ringkasan sentimen, serta saran untuk pengembangan lanjutan seperti fitur deteksi spam dan dukungan bahasa yang lebih luas.