

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG MASALAH

Dalam beberapa tahun terakhir, perkembangan teknologi telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk dalam bidang kecerdasan buatan (*Artificial Intelligence*). Salah satu cabang utama dari kecerdasan buatan adalah pembelajaran mesin (*Machine Learning*), yang memungkinkan komputer untuk belajar dari data tanpa harus diprogram secara eksplisit [1]. *Machine Learning* digunakan dalam berbagai aplikasi, mulai dari pengenalan wajah, sistem rekomendasi, hingga analisis sentimen. Dalam konteks analisis sentimen, teknik ini dapat digunakan untuk memahami dan mengklasifikasikan opini atau ulasan yang diberikan oleh pengguna terhadap suatu produk atau layanan, termasuk ulasan film [2].

Film merupakan salah satu bentuk hiburan yang memiliki pengaruh besar dalam kehidupan masyarakat. Industri perfilman terus berkembang dengan berbagai genre dan konsep yang menarik minat penonton dari berbagai kalangan. Film tidak hanya berfungsi sebagai hiburan, tetapi juga sebagai media komunikasi dan ekspresi budaya yang dapat menyampaikan pesan tertentu kepada penontonnya [3]. Dengan berkembangnya platform digital seperti IMDb, *YouTube*, maupun *Netflix*, masyarakat kini dapat dengan mudah memberikan ulasan atau pendapat mereka terhadap film yang telah mereka tonton [4]. Agar penonton terbantu melihat film yang akan datang, perlu dibuat pengingat atau cuplikan seperti suatu trailer

atau teaser akan tayang.

Ulasan film yang diberikan oleh penonton dapat memiliki dampak yang signifikan dan memberikan pengaruh besar bagi calon penonton lainnya. Ulasan yang bersifat positif dapat meningkatkan ketertarikan untuk menonton, sementara ulasan negatif bisa menurunkan minat. Bagi banyak orang, ulasan film sering dijadikan bahan pertimbangan sebelum memutuskan apakah sebuah film layak ditonton atau tidak. Oleh karena itu, penting untuk menganalisis ulasan-ulasan ini secara sistematis agar sentimen yang terkandung di dalamnya dapat dipahami dengan lebih jelas dan membantu calon penonton dalam membuat keputusan [5].

Terdapat berbagai platform yang dapat menampilkan trailer atau teaser digunakan oleh penonton untuk melihat ulasan dan penilaian terhadap sebuah film, antara lain IMDb (*Internet Movie Database*), *YouTube*, dan *Netflix*. Meskipun masing-masing platform memiliki keunikan tersendiri, IMDb menjadi salah satu pilihan yang paling relevan dalam konteks analisis ulasan dan rating, khususnya karena menyediakan data yang terstruktur serta mencakup penilaian dari pengguna umum maupun kritikus profesional. Hal ini menjadikan IMDb sebagai sumber yang kredibel dan banyak digunakan dalam penelitian terkait analisis sentimen atau evaluasi kualitas film. Berikut perbandingan platform sebagai ulasan film :

Tabel 1.1 Perbandingan Platform Sebagai Ulasan Film

Aspek	IMDb	YouTube	Netflix
Jenis Informasi	Data film formal dan rating	Ulasan, opini, dan reaksi video	Sinopsis dan trailer resmi
Rating Standar	Ya (skala 1–10)	Tidak (tidak ada sistem rating resmi)	Ya (bintang & rekomendasi internal)
Review Pengguna	Ulasan tertulis dalam bentuk text	Review dalam bentuk reaksi video	Tidak tersedia review pengguna terbuka
Trailer	Ada trailer dan cuplikan	Banyak trailer dan video terkait	Ada preview dan trailer resmi

Digunakan oleh penonton awam	Populer	Sangat populer	Sangat populer tapi fokus streaming
Digunakan oleh kritikus atau profesional	Sering digunakan untuk referensi dan penelitian	Beberapa Konten kreator profesional	Jarang untuk referensi ulasan

Sumber : IMDb (imdb.com), YouTube (youtube.com), Netflix (netflix.com).

IMDb dipilih sebagai sumber utama dalam penelitian ini untuk menentukan urutan film Indonesia terbaik berdasarkan rating pengguna. Platform ini memiliki sistem penilaian yang terpercaya serta menyediakan daftar peringkat film yang dapat dijadikan acuan objektif dalam pemilihan film yang akan dianalisis. Setelah daftar film ditentukan melalui IMDb, proses pengumpulan data ulasan dilakukan dengan memanfaatkan platform YouTube. YouTube menjadi sumber yang relevan karena menyediakan berbagai komentar dari pengguna yang secara langsung menanggapi cuplikan, trailer, atau ulasan film tersebut. Komentar-komentar ini dikumpulkan melalui metode web scraping untuk mendapatkan data teks berupa opini yang mencerminkan sentimen positif maupun negatif dari penonton terhadap masing-masing film. Dengan kombinasi pemilihan film melalui IMDb dan pengambilan data ulasan dari YouTube, penelitian ini memperoleh kumpulan data yang kaya, variatif, dan sesuai untuk dianalisis menggunakan metode klasifikasi sentimen berbasis Naïve Bayes.

Berdasarkan tabel perbandingan dari 3 platform terkenal, hanya IMDb memberikan variasi data yang kaya dan natural untuk keperluan pelatihan model klasifikasi sentimen dengan penggunaan rating. Selain itu, struktur data yang konsisten meliputi judul ulasan, isi, menggunakan rating numerik (1–10), dan nama pengguna memudahkan proses pengumpulan data dan pelabelan otomatis, misalnya dengan mengkategorikan ulasan berdasarkan ambang batas rating tertentu.

Karakteristik ini menjadikan IMDb sebagai sumber data yang sangat sesuai untuk penelitian klasifikasi sentimen.

Sebagai salah satu situs referensi film paling populer di dunia, IMDb (*Internet Movie Database*) menyusun daftar *Top Movies*, yaitu kumpulan film dengan peringkat tertinggi berdasarkan penilaian jutaan pengguna dari berbagai negara. IMDb memberikan kriteria dalam skala 1,0 hingga 10,0, dengan kategori sebagai berikut: 9,0 – 10 dinilai luar biasa (masterpiece), 8,0 – 8,9 sangat bagus (sangat direkomendasikan), 7,0 – 7,9 bagus (layak ditonton), 6,0 – 6,9 cukup (biasa saja), 5,0 – 5,9 rata-rata (ada kekurangan), 4,0 – 4,9 di bawah rata-rata (kurang baik), 3,0 – 3,9 buruk (banyak masalah) dan 1,0 – 2,9 sangat buruk (sangat tidak direkomendasikan). Kriteria ini menjadi acuan bagi penonton untuk mengetahui film-film dengan kualitas dan apresiasi tertinggi. Berikut ini ada 5 film teratas yang memiliki rating tertinggi berdasarkan tahun :

Tabel 1.2 Top 5 Indonesia di IMDb

Peringkat	Judul Film	Tahun	Rating IMDb
1	Molulo 2	2022	9.7
2	Sore Istri dari Masa Depan	2025	9.4
3	Nariti Romansa Danau Toba	2022	9.4
4	Jomblo Fi Sabilillah	2023	9.2
5	Detektif Jaga Jarak	2022	9.2

Sumber : IMDb Top Movies 2025

Berbagai penelitian sebelumnya telah dilakukan dalam upaya mengklasifikasikan sentimen ulasan film menggunakan metode *Naïve Bayes* yang menunjukkan bahwa analisis sentimen ulasan film yang dibuat menggunakan metode *Naïve Bayes* mampu mengenali apakah sebuah ulasan bersifat positif atau negatif dengan cukup baik. Setelah diuji, metode ini menunjukkan hasil yang cukup

akurat dalam memprediksi sentimen ulasan [6]. Ulasan film biasanya berupa teks. *Naïve Bayes* sangat baik dalam mengolah data berbentuk teks, karena algoritma ini menghitung kemungkinan kata-kata muncul dalam sentimen tertentu [7]. Meskipun sederhana, metode *Naïve Bayes* mampu memberikan hasil prediksi sentimen yang cukup akurat dan dapat diandalkan untuk keperluan dasar analisis [3]. Berbeda dengan metode lain yang lebih kompleks, *Naïve Bayes* tidak membutuhkan komputer dengan spesifikasi tinggi, sehingga bisa digunakan oleh banyak orang dengan alat yang sederhana [8]. *Naïve Bayes* mudah diterapkan dengan alat bantu seperti Python atau aplikasi analisis data lainnya, sehingga bisa cepat digunakan dalam proyek seperti analisis ulasan film [9].

Penelitian ini dilakukan untuk mengklasifikasikan sentimen ulasan film menggunakan metode *Naïve Bayes*, yang merupakan salah satu algoritma dalam *Machine Learning* yang efektif dalam klasifikasi teks. *Naïve Bayes* memiliki keunggulan dalam kecepatan dan akurasi yang cukup baik dalam menangani data teks yang bersifat tidak terstruktur, seperti ulasan film. Dengan menerapkan metode ini, ulasan film dapat diklasifikasikan ke dalam kategori positif dan negatif secara otomatis, sehingga membantu rumah produksi dan calon penonton dalam memahami opini publik terhadap suatu film.

Hasil dari penelitian ini dapat memberikan manfaat bagi rumah produksi film dalam mengevaluasi kualitas film yang mereka hasilkan berdasarkan opini publik. Selain itu, masyarakat juga dapat terbantu dengan adanya sistem klasifikasi sentimen ini, karena dapat memberikan gambaran umum tentang bagaimana sebuah film diterima oleh penonton secara luas. Dengan demikian, penelitian ini tidak

hanya memberikan kontribusi dalam bidang teknologi dan analisis data, tetapi juga dalam industri perfilman serta pengalaman menonton masyarakat.

Berdasarkan latar belakang diatas maka peneliti tertarik melakukan penelitian dan akan dirangkum dalam laporan tugas akhir yang berjudul “**Klasifikasi Sentimen Ulasan Film Menggunakan *Naïve Bayes***”, yang bertujuan untuk memberikan kontribusi dalam bidang analisis sentimen serta membantu pemangku kepentingan dalam industri perfilman dalam memahami opini publik terhadap sebuah film.

1.2 PERUMUSAN MASALAH

Berdasarkan uraian diatas maka rumusan masalah dalam penelitian penelitian ini adalah :

1. Membandingkan tingkat animo pengunjung platform teaser penyedia jasa film streaming internet
2. Mengapa pengunjung teaser film di IMDb rendah.
3. Bagaimana teaser film dapat diterapkan menggunakan metode *Naïve Bayes* menggunakan metrik akurasi, presisi, recall dan F1-Score.
4. Bagaimana penonton memperoleh akurasi akurat dalam menentukan data data film yang akan tayang berdasarkan tingkat kepuasan di IMDb

1.3 BATASAN MASALAH

Untuk menghindari terjadinya pembahasan di luar dari tema penelitian, maka penulis melakukan batasan masalahnya mencakup:

1. *Dataset* yang digunakan dalam penelitian ini hanya mencakup *dataset*

IMDb ulasan berbahasa Indonesia.

2. Film yang akan di ulas adalah 5 rating tertinggi pada IMDb
3. Penelitian ini akan menerapkan metode *Naïve Bayes* sebagai algoritma utama dalam klasifikasi sentimen ulasan film.
4. Sumber dataset yang digunakan dalam penelitian adalah komentar masing-masing film pada YouTube.

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah:

1. Menerapkan model klasifikasi sentimen ulasan film menggunakan metode *Naïve Bayes*.
2. Menganalisis performa model dengan mengukur tingkat akurasi, presisi, dan recall dalam mengklasifikasikan sentimen ulasan film.
3. Menyediakan hasil klasifikasi sentimen yang dapat digunakan calon penonton untuk memahami opini publik terhadap suatu film yang akan ditonton.

1.4.2 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Bagi Rumah Produksi Film: Membantu dalam mengevaluasi tanggapan penonton terhadap film yang mereka produksi, sehingga dapat menjadi referensi dalam pengembangan film berikutnya.
2. Bagi Masyarakat: Memberikan gambaran umum mengenai kualitas

sebuah film berdasarkan ulasan penonton, sehingga dapat membantu dalam pengambilan keputusan sebelum menonton.

3. Bagi Akademisi dan Peneliti: Menjadi referensi dalam penelitian lebih lanjut mengenai penerapan *Machine Learning* dalam analisis sentimen dan pemrosesan bahasa.

1.5 SISTEMATIKA PENULISAN

Penulisan ini disajikan dengan sistematika sebagai berikut :

BAB I : PENDAHULUAN

Bab ini membahas latar belakang penelitian, yaitu perkembangan teknologi kecerdasan buatan, khususnya *Machine Learning* dalam analisis sentimen ulasan film. Selain itu, diuraikan rumusan masalah yang menjadi fokus penelitian, batasan masalah untuk menjaga ruang lingkup penelitian, tujuan penelitian untuk mengklasifikasikan sentimen ulasan film menggunakan metode *Naïve Bayes*, serta manfaat penelitian bagi industri perfilman dan calon penonton. Bagian ini juga mencakup sistematika penulisan sebagai gambaran umum isi laporan.

BAB II : LANDASAN TEORI

Bab ini berisi teori-teori yang menjadi dasar penelitian. Dimulai dengan pembahasan tentang kecerdasan buatan dan *Machine Learning*, konsep analisis sentimen, serta ulasan mengenai

algoritma *Naïve Bayes* dalam klasifikasi teks. Dijelaskan kajian pustaka terhadap penelitian-penelitian terdahulu yang relevan dengan klasifikasi sentimen.

BAB III : METODOLOGI PENELITIAN

Bab ini menjelaskan pendekatan dan tahapan yang digunakan dalam penelitian. Kerangka penelitian memuat alur kerja dari pengumpulan data ulasan film, praproses data, penerapan algoritma *Naïve Bayes*, hingga evaluasi model. Dijelaskan pula alat bantu penelitian yang digunakan, baik perangkat keras maupun perangkat lunak, serta langkah teknis dari preprocessing hingga pengujian akurasi model.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dari eksperimen klasifikasi sentimen ulasan film. Disampaikan deskripsi dataset yang digunakan, implementasi algoritma *Naïve Bayes*, hasil evaluasi model menggunakan metrik seperti akurasi, precision, recall, dan F1-score, serta pembahasan terhadap performa model dalam mengklasifikasikan sentimen positif dan negatif.

BAB V : PENUTUP

Bab ini berisi kesimpulan penelitian berdasarkan hasil yang diperoleh serta jawaban atas rumusan masalah. Selain itu, disampaikan saran untuk penelitian selanjutnya, seperti pengembangan metode klasifikasi lain atau perluasan dataset.