

BAB I

PENDAHULUAN

1.1 Latar Belakang

Seiring perkembangan teknologi informasi dan komunikasi telah membuka akses yang luas bagi masyarakat untuk memberikan opini dan tanggapan terhadap berbagai produk hiburan, termasuk film. Platform seperti IMDb, Rotten Tomatoes dan berbagai media sosial lainnya telah menjadi sarana utama bagi penonton untuk mengekspresikan pendapat mereka tentang sebuah film, baik dalam bentuk ulasan singkat maupun panjang. Ulasan ini tidak hanya menjadi sumber informasi bagi calon penonton, tetapi juga berperan penting bagi para produser dan distributor dalam memahami bagaimana film mereka diterima oleh masyarakat.

IMDb (*Internet Movie Database*) dikenal sebagai salah satu platform utama yang menyajikan informasi dan ulasan film dari seluruh dunia, di mana penonton dapat memberikan rating, *review*, dan komentar terhadap berbagai film. Salah satu film yang banyak menarik perhatian publik karena mengangkat tema bernuansa agama Islam adalah film *The Lady Of Heaven*. Sebagai film yang mengangkat film dengan kisah hidup Fatimah dari seorang putri kesayangan Rasullah Muhammad menarik minat penonton dari berbagai kalangan, sehingga menciptakan volume *review* yang sangat besar di IMDb. Ulasan-ulasan untuk film ini tidak hanya mencerminkan pandangan penonton terhadap film, tetapi juga memberikan indikasi tentang elemen-elemen yang disukai atau tidak disukai, seperti alur cerita, karakter, efek visual, dan aspek teknis lainnya.

Namun, karena banyaknya jumlah ulasan menjadi tantangan besar bagi studio, produser, bahkan peneliti untuk memahami opini penonton secara keseluruhan hanya dengan membaca satu per satu opini tersebut. Oleh karena itu, analisis sentimen menjadi solusi efektif untuk mengelompokkan dan memahami pola sentimen penonton, apakah mereka memberikan tanggapan positif, negatif, atau netral terhadap film *The Lady Of Heaven*. Analisis sentimen ini memungkinkan pemetaan yang lebih sistematis terhadap opini penonton, sehingga membantu pihak terkait dalam mengevaluasi film serta memahami respons penonton.

Metode C4.5 merupakan salah satu algoritma dalam *decision tree* yang dapat digunakan untuk analisis sentimen, dengan tujuan mengklasifikasikan ulasan berdasarkan sentimen yang terkandung di dalamnya. Algoritma ini bekerja dengan membangun pohon keputusan berdasarkan atribut yang terdapat dalam data dan memberikan hasil klasifikasi yang terstruktur juga sangat mudah dipahami. Pada penelitian ini, algoritma C4.5 digunakan pada dataset ulasan film *The Lady Of Heaven* di IMDb untuk mengelompokkan sentimen penonton menjadi tiga kategori yaitu positif, negatif atau netral. Pemilihan algoritma ini didasarkan pada keunggulan algoritma *decision tree* C4.5 dalam menangani data kompleks, serta kemampuannya dalam memberikan hasil yang akurat dalam klasifikasi.

Amalia Anjani, Mochammad Fuad, dan Bagus Utomo (Arifiyanti et al., 2022) telah melakukan penelitian tentang “Analisis Sentimen Ulasan Pengunjung Objek Wisata Gunung Bromo pada Situs Tripadvisor”. Penelitian ini dilakukan untuk memberikan gambaran umum mengenai penilaian dari para pengunjung obyek wisata Gunung Bromo melalui perbandingan antara tiga algoritma yaitu Naïve Bayes, *Logistic Regression*, dan *Decision Tree* menggunakan data ulasan TripAdvisor sebanyak 500 baris data. Dari ketiga algoritma pengklasifikasi yang diuji dalam penelitian ini, algoritma *Decision Tree* mencapai tingkat akurasi model klasifikasi tertinggi yaitu 91%, diikuti oleh Naïve Bayes dan *Logistic Regression* yang masing-masing sebesar 88%. Nilai *precision*, *recall*, dan f1-measure untuk *Decision Tree* masing-masing sebesar 0,95 , 0,62 , dan 0,68. Dengan demikian, dapat disimpulkan bahwa nilai akurasi yang lebih tinggi diperoleh dengan menggunakan metode *decision tree*.

Berdasarkan analisa dan latar belakang diatas, maka penulis berminat melakukan penelitian tentang **“Analisis Sentimen Review Film Pada Platform IMDb”**. Diharapkan melalui analisis sentimen pada ulasan film di IMDb dapat menyajikan pemahaman yang lebih mendalam mengenai pandangan penonton tentang film tersebut, serta membantu IMDb dalam menyempurnakan rekomendasi film mereka. Untuk menghasilkan tingkat akurasi yang tinggi menggunakan metode *decision tree* C4.5 dilihat dari paparan penelitian terdahulu.

1.2 Rumusan Masalah

Berdasarkan pemaparan latar belakang di atas, maka dapat dirumuskan permasalahan sebagai berikut :

1. Bagaimana cara menerapkan algoritma C4.5 untuk mengklasifikasikan sentimen publik terhadap film *The Lady Of Heaven* pada platform IMDb menjadi kategori positif, negatif dan netral?
2. Bagaimana hasil keakuratan metode *decision tree* C4.5 yang diterapkan dalam analisis sentimen pada opini publik terhadap film *The Lady Of Heaven* dalam platform IMDb?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan pada latar belakang di atas, penelitian ini bertujuan untuk mencapai beberapa sasaran sebagai berikut :

1. Menerapkan algoritma *decision tree* untuk menganalisis sentimen positif, negatif dan netral pada platform IMDb terkait film *The Lady of Heaven* (2021) melalui proses *text mining*.
2. Menentukan tingkat akurasi dari metode algoritma *decision tree* dalam menganalisis film *The Lady of Heaven* (2021).

1.4 Batasan Masalah

Penetapan batasan masalah dilakukan untuk memastikan bahwa penelitian ini dapat mencapai tujuan yang telah ditetapkan. Adapun batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Penelitian ini hanya berfokus pada satu film yaitu *The Lady of Heaven* (2021).
2. Data yang diambil menggunakan komentar berbahasa inggris pada platform IMDb.
3. Data yang digunakan berjumlah 1512.
4. Data yang dikumpulkan berasal dari periode Desember 2021 hingga Juni 2022.
5. Data masukan terdiri dari tweet yang berisi komentar penonton, tanpa menyertakan gambar atau video.

6. Hasil analisis sentimen dikategorikan kedalam 3 kelas yaitu positif, negatif, dan netral.
7. Metode penelitian yang diterapkan dalam penelitian ini adalah algoritma *decision tree C4.5*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat. Berikut adalah manfaat yang diharapkan dari penelitian ini :

1. Memudahkan proses pengklasifikasian komentar pada platform IMDb menggunakan bahasa inggris.
2. Mengetahui tingkat akurasi metode algoritma *decision tree* dalam pengklasifikasian *text mining*.
3. Dapat dijadikan referensi dalam pendistribusian film selanjutnya, sehingga meningkatkan kualitas karya mereka.
4. Dapat digunakan sebagai acuan untuk penelitian lanjutan di masa mendatang.
5. Membantu penonton memahami ulasan atau persepsi umum terhadap film *The Lady of Heaven*.