

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Dengan pesatnya perkembangan teknologi komunikasi dan informasi pada zaman sekarang, film menjadi salah satu konten media yang paling banyak diminati warganet dengan beragam jenis genre dan judul; bahkan warganet juga dapat menontonnya dengan mudah secara online pada situs penyedia jasa streaming resmi berbayar seperti Netflix, Hulu, Apple TV+, Amazon Prime Video, Disney+, dll. Warganet yang sudah menonton sebuah film dapat memberikan ulasan positif dan negatif pada situs IMDB (*Internet Movie Database*) yang menjadi database online pelbagai informasi tentang film [1]. Ulasan positif dan negatif sebuah film pada situs IMDB (*Internet Movie Database*) dapat menjadi referensi calon penonton yang ingin menonton atau membeli sebuah film.

Pada bulan Juli-September 2023 film *Oppenheimer* menjadi trending topik pada Google Trends Worldwide [2] dan pada tahun 2023 film *Oppenheimer* menjadi trending topik pada Google's Year in Search Global diurutan ke-2 dalam kategori Movies [3]. Penelitian ini menggunakan dataset ulasan film *Oppenheimer* pada situs IMDB (*Internet Movie Database*) yang diunduh dari situs Kaggle sebanyak 3857 baris data [4].

Analisis sentimen merupakan proses menganalisa data text yang memiliki sentimen positif, negatif atau netral [5]. Penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) dalam melakukan analisis sentimen positif dan negatif ulasan film *Oppenheimer* pada situs IMDB (*Internet Movie Database*) yang merupakan salah satu metode *machine learning* yang digunakan untuk mengklasifikasi data menggunakan garis pembatas *hyperplane* dalam bentuk vector [6][7][8].

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah peneliti uraikan, maka rumusan masalah pada penelitian ini sebagai berikut:

14. Bagaimana menerapkan algoritma *Support Vector Machine* (SVM) dalam melakukan analisis sentimen ulasan film Oppenheimer pada situs IMDB (*Internet Movie Database*).
15. Bagaimana hasil akurasi algoritma *Support Vector Machine* (SVM) dalam melakukan analisis sentimen ulasan film Oppenheimer pada situs IMDB (*Internet Movie Database*).

1.3 Batasan Masalah

Batasan masalah pada penelitian ini sebagai berikut:

1. Menggunakan algoritma *Support Vector Machine* (SVM) kernel linear dalam melakukan analisis sentimen positif dan negatif.
2. Menggunakan teknik TF-IDF (*Term Frequency-Inverse Document Frequency*) dalam melakukan perhitungan persentase kemunculan dan bobot kata.
3. Analisis sentimen ulasan film Oppenheimer dibagi menjadi dua kategori, yaitu: positif dan negatif.
4. Menggunakan 3857 baris data ulasan film Oppenheimer pada situs IMDB (*Internet Movie Database*) yang datasetnya diunduh dari situs Kaggle.
5. Dataset pada penelitian ini dibagi menjadi 80% data training dan 20% data testing.
6. Pelabelan sentimen positif dan negatif pada ulasan film Oppenheimer dilakukan secara manual oleh peneliti.
7. Data pada ulasan film Oppenheimer menggunakan ulasan berbahasa Inggris.
8. Menggunakan teknik *K-Fold Cross Validation* dan *Confusion Matrix* dalam menguji dan mengevaluasi algoritma *Support Vector Machine* (SVM).

1.4 Tujuan Penelitian

Tujuan peneliti dalam melakukan penelitian ini sebagai berikut:

1. Sebagai syarat kelulusan pada jenjang Strata (S1) 1 Universitas AMIKOM Yogyakarta.
2. Dapat menerapkan algoritma *Support Vector Machine* (SVM) dalam melakukan analisis sentimen positif dan negatif ulasan film Oppenheimer pada situs IMDB (*Internet Movie Database*).
3. Dapat menguji dan mengevaluasi tingkat akurasi algoritma *Support Vector Machine* (SVM) dalam melakukan analisis sentimen positif dan negatif ulasan film Oppenheimer pada situs IMDB (*Internet Movie Database*) menggunakan dataset yang sudah diberi label sentimen positif dan negatif secara manual.

1.5 Manfaat Penelitian

1. Memberikan informasi hasil analisis sentimen positif dan negatif ulasan film Oppenheimer pada situs IMDB (*Internet Movie Database*) menggunakan algoritma *Support Vector Machine* (SVM).
2. Memberikan informasi tata cara penerapan algoritma *Support Vector Machine* (SVM) dalam melakukan analisis sentimen positif dan negatif ulasan film Oppenheimer pada situs IMDB (*Internet Movie Database*).

1.6 Sistematika Penulisan

Penyusunan dan penulisan dalam penelitian ini menggunakan sistematika sebagai berikut:

BAB I: PENDAHULUAN

Pada bab ini membahas tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II: TINJAUAN PUSTAKA

Pada bab ini membahas tentang studi literatur dan dasar teori dari pelbagai sumber tentang *Data Mining*, *Text Mining*, *Text Preprocessing*, *TF-IDF* (*Term Frequency-Inverse Document*

Frequency), *Support Vector Machine (SVM)*, *K-Fold Cross Validation* dan *Confusion Matrix* serta membahas dan membandingkan tinjauan pustaka yang digunakan dalam melakukan penelitian.

BAB III: METODE PENELITIAN

Pada bab ini membahas tentang objek penelitian, alur penelitian serta alat dan bahan yang digunakan dalam melakukan penelitian.

BAB IV: HASIL DAN PEMBAHASAN

Pada bab ini membahas tentang tahap-tahap dan implementasi algoritma *Support Vector Machine (SVM)*, *Text Preprocessing*, *TF-IDF (Term Frequency-Inverse Document Frequency)* serta menguji dan mengevaluasi algoritma *Support Vector Machine (SVM)* menggunakan *K-Fold Cross Validation* dan *Confusion Matrix*.

BAB V: PENUTUP

Pada bab ini membahas tentang kesimpulan yang didapat berdasarkan hasil penelitian yang telah dilakukan dan saran kepada peneliti selanjutnya.