

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Selama beberapa dekade terakhir, internet telah menjadi kebutuhan dan gaya hidup global, termasuk di Indonesia. Perubahan gaya hidup, kebiasaan, sosial, ekonomi, dan budaya telah diubah secara signifikan oleh internet, meskipun tanpa disadari. Internet memungkinkan akses ke berbagai aspek kehidupan seperti komunikasi, bisnis, informasi, pendidikan dan hiburan seperti musik dan film[1].

Film berfungsi sebagai sarana baru untuk menyebarkan hiburan yang sudah lama ada, serta menyajikan cerita, peristiwa, musik, drama, komedi, politik dan berbagai sajian teknis lainnya kepada masyarakat umum[2]. Salah satu film Indonesia yang menarik untuk diteliti adalah "Dirty Vote", yang mengangkat isu politik yang kontroversial dan memicu berbagai komentar serta diskusi di YouTube.

Film dokumenter "Dirty Vote" merupakan sebuah investigasi mendalam mengenai dugaan kecurangan dalam pelaksanaan Pemilu 2024 di Indonesia. Film ini menyoroti berbagai praktik politik yang dianggap tidak sehat, seperti penggunaan sumber daya negara untuk kepentingan kampanye, manipulasi data pemilih, hingga upaya-upaya untuk memengaruhi opini publik. Dengan mengupas tuntas berbagai kasus dan wawancara dengan para ahli, "Dirty Vote" berusaha membuka mata publik tentang pentingnya menjaga integritas pemilu dan bahaya dari praktik-praktik kotor dalam politik.

Analisis sentimen, juga dikenal sebagai penambangan opini, yaitu cabang ilmu yang bertujuan untuk menganalisis pendapat dan perasaan individu terhadap berbagai entitas. Entitas-entitas ini bisa berupa produk, layanan, organisasi, individu, isu, peristiwa, atau topik tertentu[3]. Dalam konteks film "Dirty Vote", analisis sentimen dapat memberikan wawasan tentang bagaimana penonton menilai film tersebut, baik secara positif maupun negatif. Informasi ini penting bagi

pembuat film, distributor, dan peneliti untuk memahami preferensi penonton film Indonesia.

Ada berbagai metode yang dapat digunakan untuk analisis sentimen, seperti Complement Naive Bayes (CNB) dan Support Vector Machine (SVM). Kedua metode ini terbukti efektif dalam mengklasifikasikan teks ke dalam kategori sentimen positif, negatif, atau netral. Namun, performa metode tersebut sangat tergantung pada kualitas data yang digunakan. Oleh karena itu, seleksi fitur diperlukan untuk memilih kata-kata yang paling relevan dengan sentimen terhadap film "Dirty Vote".

Dalam penelitian ini, metode seleksi fitur Chi-Square dan Information Gain (IG) akan digunakan untuk memilih kata-kata yang paling informatif dari komentar YouTube tentang film "Dirty Vote". Kata-kata yang terpilih akan menjadi input untuk model Complement Naive Bayes dan SVM guna mengklasifikasikan sentimen komentar. Diharapkan penelitian ini dapat memberikan wawasan yang berharga tentang persepsi masyarakat terhadap film "Dirty Vote" serta meningkatkan kinerja analisis sentimen dengan menggunakan metode CNB dan SVM. Berdasarkan uraian di atas maka penulis tertarik untuk melakukan penelitian dengan studi kasus film "Dirty Vote" yang berjudul **"Perbandingan Analisis Sentimen Menggunakan Complement Naive Bayes dan Support Vector Machine"**

## **1.2 Rumusan Masalah**

Berdasar latar belakang di atas, dapat dirumuskan sebuah permasalahan yaitu :

1. Apakah metode seleksi fitur Chi-Square dan IG dapat meningkatkan kinerja model Naive Bayes dan SVM dalam mengklasifikasikan sentimen komentar YouTube terhadap film "Dirty Vote"?
2. Model klasifikasi sentimen mana yang lebih efektif untuk mengklasifikasikan komentar YouTube terhadap film "Dirty Vote", Naive Bayes atau SVM?

### 1.3 Batasan Masalah

Berdasarkan rumusan masalah penelitian ini memiliki beberapa batasan yaitu :

1. Data yang digunakan dalam penelitian ini hanya berasal dari komentar YouTube terhadap film "Dirty Vote". Data yang digunakan adalah komentar yang diambil sampai 22 Maret 2024.
2. Penelitian ini hanya menggunakan dua metode klasifikasi sentimen, yaitu Naive Bayes dan SVM. Metode klasifikasi sentimen lainnya tidak dianalisis dalam penelitian ini.
3. Penelitian ini hanya menggunakan 2 metode seleksi fitur, yaitu Chi-Square dan Information Gain.

### 1.4 Tujuan Penelitian

Berdasarkan uraian yang telah dipaparkan, penelitian ini memiliki tujuan dan sasaran sebagai berikut:

1. Membandingkan kinerja model Naive Bayes dan SVM dalam mengklasifikasikan sentimen komentar YouTube terhadap film "Dirty Vote" dengan dan tanpa seleksi fitur Chi-Square dan IG.
2. Menentukan model klasifikasi sentimen yang paling efektif untuk mengklasifikasikan komentar YouTube terhadap film "Dirty Vote", Naive Bayes atau SVM.

### 1.5 Manfaat Penelitian

Penelitian ini diharapkan bisa memberikan manfaat sebagai berikut:

1. Secara teoritis penelitian ini diharapkan dapat memberikan kontribusi baru pada ilmu pengetahuan di bidang analisis sentimen, khususnya dalam konteks analisis sentimen terhadap film di YouTube.
2. Penelitian ini juga diharapkan dapat meningkatkan pemahaman tentang efektivitas metode seleksi fitur Chi-Square dan IG dalam meningkatkan

kinerja model klasifikasi sentimen.

3. Secara praktis hasil penelitian ini dapat digunakan sebagai referensi untuk penelitian lain di bidang analisis sentimen, pemrosesan bahasa alami, dan media sosial.

## **1.6 Sistematika Penulisan**

Dalam penelitian ini sistematika penulisan terdiri dari lima bab, dengan susunan sebagai berikut :

### **Bab I Pendahuluan**

Berisikan latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, Batasan masalah, dan sistematika penulisan

### **Bab II Landasan Teori**

Berisi tentang pembahasan teori-teori yang mendukung proses pembuatan Tugas Akhir ini

### **Bab III Metode Penelitian**

Berisi tentang sumber data dan tahapan yang dilakukan dalam pembuatan penelitian serta gambaran umum sistem yang akan dikerjakan

### **Bab IV Hasil dan Pembahasan**

Berisi tentang hasil penelitian dan pembahasannya

### **Bab V Penutup**

Berisi tentang kesimpulan dan saran sebagai bahan evaluasi penulis.