

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Twitter adalah suatu media sosial yang berkembang pesat saat ini, dimana penggunanya dapat memposting, dan membaca pesan suatu teks pendek maupun panjang atau yang disebut sebagai tweet. Melalui tweet tersebut pengguna twitter dapat menyampaikan pendapat, pandangan dan pemikiran mereka tentang topik tertentu. Twitter kerap dijadikan sebagai sumber data pendapat dan sentimen masyarakat untuk dianalisis, dimana data tersebut dapat digunakan untuk memahami pendapat masyarakat tentang suatu topik. Dimana tujuan utama dilakukannya analisis sentimen di Twitter adalah untuk menentukan opini dari suatu tweet, apakah bernilai positif atau negatif [1].

Data dari Statista Research Department menunjukkan per Desember 2022, pengguna Twitter di seluruh dunia mencapai lebih dari 368 juta pengguna aktif bulanan. Angka ini diperkirakan akan menurun menjadi sekitar 335 juta pada tahun 2024, mengalami penurunan sekitar lima persen dibandingkan dengan tahun 2022 [2]. Sebagai media sosial tempat berbagi pendapat yang memiliki banyak pengguna yang tersebar di seluruh dunia, Twitter setiap harinya memiliki banyak tweet yang berisi berbagai pendapat dengan beragam topik yang dibicarakan. Hal tersebut membuat banyaknya penyebaran dan pertumbuhan informasi di Twitter terus meningkat [3]. Salah satu topik yang banyak mendapat perhatian di Twitter adalah kasus-kasus kriminal, seperti pembunuhan, yang sering kali memicu berbagai tanggapan dari pengguna dengan sudut pandang beragam.

Kasus pembunuhan di Indonesia sering kali mencerminkan kompleksitas sosial, ekonomi, dan budaya dalam masyarakat. Salah satu kasus yang sempat viral adalah pembunuhan Vina yang kemudian dijadikan film yang berjudul *Vina: Sebelum 7 Hari*. Masyarakat menilai film ini tidak layak ditayangkan karena dianggap tidak etis dan tidak bermoral. Film ini mendapat kecaman dari para kritikus film dan terancam diboikot karena dianggap mengeksploitasi suatu

peristiwa. Setelah dirilis, muncul berbagai ulasan terutama pada Twitter yang memberikan pendapat negatif maupun positif. Oleh karena itu, penelitian ini difokuskan untuk menganalisis sentimen ulasan tersebut dengan tujuan untuk mengetahui apakah masyarakat menilai film ini layak ditonton atau tidak.

Menurut penelitian [4] yang membandingkan performa algoritma Naive Bayes, SVM, dan Random Forest terhadap Ibu Kota Nusantara (IKN), model dengan akurasi tertinggi adalah SVM dengan 87%, diikuti oleh Random Forest dengan akurasi 85%, dan Naive Bayes dengan akurasi 73,9%.

Dari hasil penelitian [5] yang berjudul Analisis Sentimen Aplikasi Halo BCA Di Google Play Store Menggunakan Metode Naive Bayes, Support Vector Machine Dan Random Forest. Hasil akurasi tertinggi adalah Random Forest dengan akurasi sebesar 91,28%, presisi 0,91, recall 0,91, dan F1-score 0,91. SVM memiliki akurasi sebesar 87,55%, sedangkan Naive Bayes 81,73%

Studi lain oleh [6] menghasilkan sebuah sistem analisis sentimen pada jasa ojek online menggunakan metode Naive Bayes. Pada penelitian ini menggunakan tweet mention jasa ojek online di Twitter. Hasil dari pengujian menggunakan Naive Bayes mendapatkan akurasi sebesar 80%.

Berdasarkan penelitian terdahulu telah menunjukkan bahwa ketiga model algoritma tersebut memberikan hasil yang efektif dalam analisis sentimen. Berdasarkan hal tersebut, penelitian ini akan menggunakan kinerja tiga model algoritma, yaitu Naive Bayes, Random Forest dan Support Vector Machine (SVM), dalam konteks analisis sentimen terhadap film *Vina: Sebelum 7 Hari*.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka rumusan masalah pada penelitian ini, yaitu:

1. Bagaimana sentimen pengguna Twitter terhadap film *Vina: Sebelum 7 Hari*?
2. Bagaimana akurasi algoritma Naïve Bayes, Random Forest dan SVM dalam melakukan analisis sentimen terhadap kasus film *Vina: Sebelum 7 Hari*?

## 1.3 Batasan Masalah

1. Model algoritma yang digunakan dalam klasifikasi data adalah Naïve Bayes, Random Forest dan SVM.
2. Tweet yang digunakan hanya tweet berbahasa Indonesia.
3. Pengambilan data dari Twitter menggunakan Tweet Harvest.
4. Memiliki dua label klasifikasi sentimen yaitu sentimen positif dan negatif.
5. Bahasa pemrograman yang digunakan untuk pengolahan data analisis sentimen adalah Python dengan menggunakan Google Colaboratory.
6. Bahasa pemrograman yang digunakan untuk Deployment adalah Python dengan menggunakan Visual Studio Code.

## 1.4 Tujuan Penelitian

Tujuan di buatnya penelitian ini yaitu :

1. Mengetahui sentimen pengguna Twitter terhadap film *Vina: Sebelum 7 Hari* apakah positif atau negatif.
2. Mengetahui performa algoritma Naïve Bayes, Random Forest dan SVM dalam melakukan klasifikasi tweet pada Twitter terhadap film *Vina: Sebelum 7 Hari*.

## 1.5 Manfaat Penelitian

Manfaat pada penelitian ini yaitu :

1. Memberikan gambaran mengenai sentimen pengguna Twitter terhadap film *Vina: Sebelum 7 Hari*.
2. Untuk menambah wawasan dan informasi mengenai implementasi model algoritma Naïve Bayes, Random Forest dan SVM dalam analisis sentimen bagi peneliti selanjutnya.
3. Untuk memberikan referensi dan sebagai acuan bagi pembaca dan peneliti selanjutnya dalam penerapan model algoritma Naïve Bayes, Random Forest dan SVM.

#### **1.6 Sistematika Penulisan**

Agar dapat mempermudah dalam memahami maupun membaca penelitian ini, maka peneliti membuat beberapa bab sebagai berikut :

##### **BAB I PENDAHULUAN**

Bab ini menguraikan bagaimana latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

##### **BAB II LANDASAN TEORI**

Bab ini berisi tinjauan pustaka dan dasar-dasar teori yang digunakan untuk menjadi acuan dalam pembuatan skripsi.

##### **BAB III METODE PENELITIAN**

Bab ini menguraikan mengenai penyelesaian masalah berdasarkan metodologi yang telah dipilih serta berisi proses implementasi dari metode tersebut.

##### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini berisi mengenai hasil yang didapat dari pengujian yang telah dilakukan pada penelitian ini serta evaluasi dari penelitian tersebut.

##### **BAB V PENUTUP**

Bab ini berisikan kesimpulan dan hasil akhir dari penelitian yang telah dilakukan serta berisi saran untuk perbaikan.