

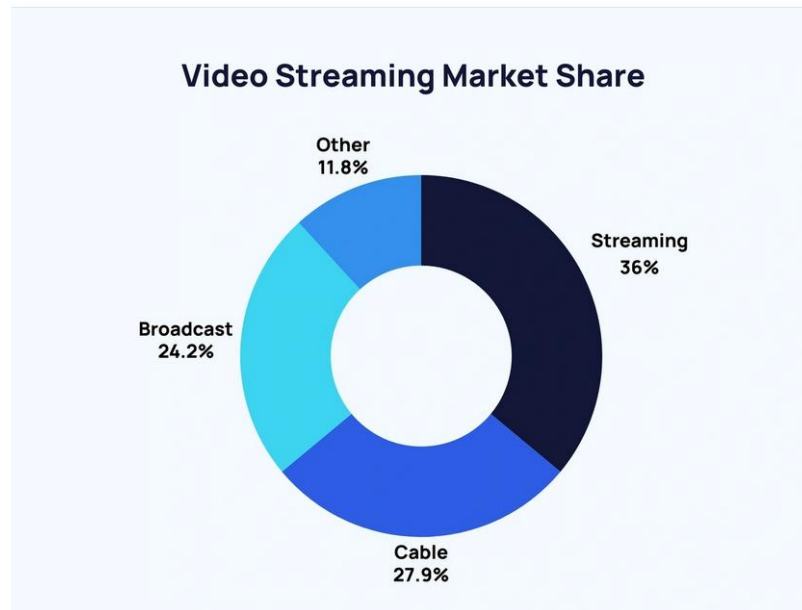
## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Evolusi *streaming video* sudah mengalihkan cara masyarakat mengonsumsi informasi dan hiburan akibat pertumbuhannya yang sangat pesat. Banyak layanan *streaming video*, termasuk *YouTube*, *Netflix*, *Disney+*, dan platform lokal, menjadi semakin populer karena perkembangan teknologi internet yang semakin cepat dan lebih terjangkau. Saat ini, masyarakat lebih menyukai kemudahan dan fleksibilitas untuk menonton konten di berbagai perangkat kapan saja dan dari lokasi mana pun. Untuk mengakomodasi minat dan preferensi pemirsa yang beragam, konten yang disajikan juga semakin beragam, meliputi dokumenter, acara TV, *film*, dan konten yang diproduksi oleh kreator individu. Fitur *live streaming* dan komentar juga mendorong lebih banyak interaksi antara produsen konten dan pemirsa, sehingga menghasilkan komunitas daring yang aktif. Selain mengubah lanskap media dan hiburan, fenomena ini menciptakan model bisnis dan peluang baru bagi konten kreator.

Menurut data statistik pada tahun 2024, industri *streaming video* saat ini bernilai lebih dari \$670 miliar. Dengan *Compound Annual Growth Rate (CAGR)* sebanyak 17,8%, industri diprediksi memperoleh \$2,49 triliun lebih pada tahun 2032. Demikian pula, *streaming* memiliki pangsa pasar yang lebih besar daripada siaran atau kabel, yang mencakup 36% dari semua tayangan TV. Berikut merupakan tampilan grafik dari pasar dari layanan tersebut:



**Gambar 1.1** Lanskap Media Digital 2024

Sumber: Nielsen

Kemudian pada data lainnya berdasarkan precedence research menunjukkan hasil analisis data bahwa industri streaming video di seluruh dunia akan mencapai USD 129,80 miliar pada tahun 2024 dan sekitar USD 865,85 miliar pada tahun 2034, berkembang pada *CAGR* yang kuat sebesar 20,90% antara tahun 2024 dan 2034. Pada tahun 2023, ukuran pasar streaming video Amerika Utara adalah USD 34,19 miliar. Menurut statistik *goodstats*, *Netflix*, *Disney+*, *Hotstar*, *Vidio*, *Viu*, *WeTV*, *iflix*, dan *Amazon Prime Video* termasuk platform yang paling sering digunakan di Indonesia. Menurut survei yang dirilis pada akhir tahun 2024, *Vidio* menjadi sebagai platform streaming dengan jumlah pelanggan berbayar terbanyak di Indonesia, yakni sebanyak 4,7 juta pengguna, dengan perkiraan 11,5 juta pelanggan berbayar layanan *streaming video* pada tahun 2023. Menurut laporan,

*Disney+ Hotstar* memiliki lebih dari 2,5 juta pengguna, diikuti *Viu* 1,5 juta, *Vidio* 1,1 juta, dan *Netflix* 850 ribu. Berdasarkan data yang disebutkan di atas, masyarakat tampaknya lebih menyukai *streaming video* untuk hiburan. Setiap *platform* bertujuan untuk menyediakan layanan terbaik di pagi hari bagi penggunanya dengan biaya berlangganan yang relatif terjangkau. Baik aspek positif maupun negatif masyarakat memiliki dampak yang besar secara signifikan oleh penggunaan *platform streaming video*. Manfaatnya meliputi akses yang mudah dan tak terbatas ke berbagai konten informasi dan hiburan, termasuk dokumenter, acara televisi, *film*, dan materi pendidikan. Selain itu, platform ini mendorong pertumbuhan orisinalitas dan kreativitas dalam pembuatan konten, sehingga memberikan kesempatan kepada produsen konten independen untuk terhubung dengan audiens di seluruh dunia.

Pengalaman menonton yang lebih individual juga disediakan melalui *streaming video*, sering kali melalui saran atau rekomendasi berdasarkan preferensi menonton dan batasan geografis yang menunjukkan karya seni dari seluruh dunia. Namun, ada dampak negatif yang harus dipertimbangkan selain kenyamanan dan kenikmatan yang diberikan. Salah satunya adalah kemungkinan gangguan dan berkurangnya produktivitas akibat akses yang terlalu mudah, yang dapat menyebabkan orang menghabiskan terlalu banyak waktu menonton untuk mengalihkan tanggung jawab utama mereka, terutama bagi mereka yang berusia rentan seperti anak-anak dan remaja. Ketergantungan pada layanan *streaming* juga dapat menyebabkan gaya hidup yang lebih menyendiri dan berkurangnya hubungan sosial secara langsung. Secara ekonomi, dominasi platform besar dapat

menyebabkan persaingan yang tidak sehat dan matinya sektor media konvensional. Cukup memprihatinkan ketika anak-anak dan remaja terpapar konten yang berisi hal-hal negatif atau tidak sesuai usia tanpa pengawasan yang tepat. Selain itu, sebagian masyarakat mungkin merasa kesulitan secara finansial untuk membayar langganan beberapa layanan streaming jika tidak bisa dikelola dengan bijak. Oleh karena itu, konsumen harus menggunakan layanan *streaming video* secara bijaksana dan secukupnya, dengan mempertimbangkan dampaknya terhadap aspek kehidupan sosial, ekonomi, dan psikologis.

Kebutuhan untuk menilai opini penonton *film* secara otomatis dan efektif dalam menghadapi pertumbuhan layanan *streaming video* merupakan latar belakang dari penelitian ini. Menganalisis sejumlah besar sentimen di platform ini menjadi tantangan sulit. Untuk mengklasifikasikan sentimen (positif, negatif, atau netral) secara akurat dari komentar penonton *film* pada layanan *streaming video*, penelitian ini bertujuan untuk melakukan pengembangan metode klasifikasi opini. Dengan menggunakan data komentar pengguna yang didapatkan dari *Google Play Store* melalui *web scraping*, studi ini menerapkan pendekatan *datamining* dan *Naive Bayes* guna mencapai tujuan ini. Teorema *Bayes*, bekerja dengan memperkirakan kemungkinan suatu peristiwa berdasarkan data sebelumnya (Al-Fajr Ramadhani et al., 2024). Semua tahapan analisis, seperti pra-proses, pelabelan data, klasifikasi, serta penilaian model, metrik akurasi, presisi, *recall*, dan skor-F1, dengan menggunakan *Google Colab (Colaboratory)* untuk menjalankan kode *Python*, Sebuah model klasifikasi opini algoritma *Naive Bayes* dengan kinerja yang dapat diukur dan pemahaman yang lebih baik tentang distribusi sentimen penonton

terhadap berbagai *film* pada *platform streaming video* yang tercipta berdasarkan dari riset. Temuan riset ini ditujukan untuk mendukung penonton dan penyedia layanan streaming video lebih baik dalam memahami dan memanfaatkan opini publik.

Penelitian sebelumnya membandingkan naive bayes dan SVM dalam menganalisis opini pengguna aplikasi *Vidio*, hasil temuan yaitu *KNN* berkinerja lebih baik daripada *Naive Bayes* di sejumlah area yang signifikan. Akurasi *KNN* adalah 74,92%, sedangkan *Naive Bayes* adalah 71,32%. *KNN* juga berkinerja lebih tinggi dalam hal presisi, dengan nilai 76,52%, dibandingkan dengan 71,61% untuk *Naive Bayes* (Pratmanto et al., 2024) Penelitian lainnya kalsifikasi *naive bayes* digunakan untuk mengklasifikasikan kenaikan pangkat pada dinas ketenaga kerjaan, hasil studi mengklasifikasikan data kelayakan calon karyawan untuk kenaikan jabatan dengan efektif dan akurat (Yoseva Simanjuntak & Septian Salomo Simatupang, 2022).

Penulis tertarik untuk melakukan penelitian dalam bentuk analisis data histori data mining dengan metode *naive bayes*, berdasarkan latar belakang dan uraian permasalahan tersebut. Penulis meneliti sebagai tugas akhir skripsi dengan judul **“Klasifikasi Opini Penonton Film Pada *Platform Streaming Video* dengan algoritma *Naïve Bayes*”**.

## **1.2 Identifikasi Masalah**

Permasalahan berikut diidentifikasi berlandaskan pembahasan latar belakang, dan mencerminkan maksud serta tujuan penelitian ini:

1. Pada layanan *streaming video* seperti *Netflix* dan *Disney+*, jumlah komentar dan opini pengguna terus bertambah, tetapi saat ini belum ada metode otomatis yang efektif untuk mengkategorikan sentimen komentar ini secara akurat.
2. Seberapa baik kinerja algoritma *Naïve Bayes* dalam mengidentifikasi sentimen komentar pengguna pada *platform streaming video* dan bagaimana penyedia layanan dapat menggunakan hasil klasifikasinya untuk memberikan wawasan, masih belum diketahui dengan jelas.

### 1.3 Batasan Masalah

Riset ini dibatasi pada permasalahan berikut:

1. Berdasarkan komentar pengguna dari *Google Play Store* diambil melalui *web scraping*, studi ini hanya berfokus pada klasifikasi sikap audiens pada layanan *streaming video Netflix* dan *Disney+*.
2. Algoritma *Gaussian Naïve Bayes* merupakan algoritma yang digunakan dalam proses klasifikasi opini, tanpa membandingkan dengan algoritma klasifikasi lain seperti *SVM* atau *Decision Tree*.

### 1.4 Rumusan Masalah

Rumusan masalah dalam studi ini adalah sebagai berikut:

1. Bagaimana opini penonton film pada layanan *streaming video Netflix* dan *Disney+* dapat dikategorikan menjadi sentiment positif, negative, dan netral dengan penerapan *Naïve Bayes*?
2. Seberapa baik algoritma *Naïve Bayes* mengkategorikan komentar pengguna menggunakan kriteria penilaian akurasi, presisi, , dan skor F1?

## 1.5 Tujuan penelitian

Adapun tujuan dari telaah ini diuraikan sebagai:

1. Guna mengkategorikan opini penonton film terhadap layanan *streaming video Netflix dan Disney+* menjadi sentimen positif, negative dan netral melalui pendekatan *Naïve Bayes*.
2. Untuk menggunakan metrix evaluasi seperti akurasi, presisi, *recall*, serta skor F1 guna menilai hasil penenrapan *Naïve Bayes* dalam mengklasifikasi sentiment penonton.

## 1.6 Manfaat Penelitian

### 1.6.1 Manfaat Teoritis

1. Menambah ilmu dan pengetahuan peneliti, penyedia layanan *streaming video*, dan penonton untuk memahami serta memanfaatkan opini publik (berkaitan dengan klasifikasi opini penonton *film* pada *platform streaming video* menggunakan algoritma *naïve bayes*),
2. Menjadi sumber referensi dan inspirasi untuk penelitian lebih lanjut untuk menganalisis isu ini (berkaitan dengan klasifikasi opini penonton *film* pada situs *streaming video* menggunakan algoritma *naïve bayes*)
3. Meningkatkan kemampuan penulis untuk mengklasifikasikan data histori *datamining* secara akurat melalui *naïve bayes* sesuai temuan telaah ini.

### 1.6.2 Manfaat Praktis

1. Bagi Peneliti

- a. Memperdalam pemahaman untuk menganalisis sesuai dengan subjek yang dilakukan oleh peneliti untuk mengklasifikasi opini secara akurat dalam membagi sentimen dari komentar penonton *film* di *platform streaming video* menjadi sentiment, positif, negatif, dan netral.
- b. Melalui kinerja yang terukur serta pemahaman menyeluruh tentang distribusi opini penonton terhadap berbagai *film* di *platform streaming video*, penulis akan menerapkan pengetahuan dan konsep yang diperoleh selama penelitian untuk memprediksikan kemungkinan suatu peristiwa berdasarkan data sebelumnya.

## 2. Bagi Penyedia Layanan *Streaming Video*

Berdasarkan opini penonton dapat menunjang popularitas *platform streaming video* dengan sentimen positif, negatif, netral. Agar *platform* dapat meningkatkan layanan yang lebih berkualitas di masa mendatang.

## 3. Bagi Pengguna (Penonton)

- a. Berdasarkan opini pengguna tentang *platform* dan *film* yang disediakan di layanan *streaming video*, pengguna dapat memilih tayangan dan rekomendasi yang lebih menarik.
- b. Dalam hal ini, orang tua dapat mengevaluasi tayangan yang lebih baik untuk anak-anak, sehingga dapat berdampak positif pada tumbuh kembang anak.