



ANALISIS SENTIMEN PENGGUNA APLIKASI *NETFLIX* PADA ULASAN *GOOGLE PLAYSTORE* MENGGUNAKAN METODE *NAÏVE BAYES*

Tommy Suhendra¹, Bunga Intan², A Taqwa Martadinata³

Fakultas Ilmu Teknik, Universitas Bina Insan, Lubuklinggau, Indonesia

e-mail: 2002030054@mhs.univbinainsan.ac.id¹, bungaintan@univbinainsan.ac.id²,
taqwa@univbinainsan.ac.id³

Abstrak

Salah satu penyedia layanan berlangganan video on demand yang paling populer untuk Tv shows dan film adalah *Netflix*. Kemajuan dan penyebaran teknologi, informasi, dan komunikasi telah mengubah cara, pola, dan kualitas konsumsi media. Salah satu perubahan yang dapat di rasakan adalah munculnya teknologi yang memungkinkan pemirsa untuk melakukan *binge watching* selama berjam-jam berturut-turut langsung melalui *streaming online*. *Netflix* adalah layanan *streaming* yang menawarkan berbagai acara TV pemenang penghargaan, film, *anime*, dokumenter, dan banyak lagi di ribuan perangkat yang terhubung ke Internet. Kamu bisa menonton sepuasnya, kapan pun kamu mau—semuanya dengan satu harga bulanan yang murah. Interaksi antara penonton terjadi melalui fitur komentar, yang dapat bersifat positif atau negatif. Penelitian ini bertujuan untuk mengetahui sentimen terhadap aplikasi *Netflix* di *google playstore* menggunakan *Machine Learning* dengan algoritma *Naïve Bayes*. Data yang di ambil adalah 7000 data ulasan yang kemudian dilakukan *preprocessing* data sehingga tersisa menjadi 6999 data yang bisa diolah. *Labeling* data dilakukan otomatis berdasarkan *score rating* ulasan. Hasil *labeling* data dibagi menjadi 2 kelas yaitu kelas positif sebanyak 1676, dan kelas negatif sebanyak 5324. Klasifikasi dan evaluasi dilakukan menggunakan metode *Naïve Bayes*. Berdasarkan nilai perbandingan data *training* dan *testing* sebesar 8:2, memperoleh hasil tingkat akurasi sebesar 84% kemudian presisi 87% dan *recall* 92%.

Kata Kunci— *Netflix*; Analisis Sentimen; *Google Playstore*; *Machine Learning*; *Naïve bayes*

Abstract

Abstract— One of the most popular video on demand subscription service providers for TV shows and films is *Netflix*. The progress and spread of technology, information and communication have changed the ways, patterns and quality of media consumption. One change that can be felt is the emergence of technology that allows viewers to binge watch for hours in a row directly via online streaming. *Netflix* is a streaming service that offers a variety of award-winning TV shows, movies, anime, documentaries, and more on thousands of Internet-connected devices. You can watch as much as you want, whenever you want—all for one low monthly price. Interaction between viewers occurs through the comments feature, which can be positive or negative. This research aims to determine sentiment towards the *Netflix* application on the *Google Play Store* using *Machine Learning* with the *Naïve Bayes* algorithm. The data taken was 7000 review data which was then preprocessed so that the remaining data was 6999 data that could be processed. Data labeling is done automatically based on review rating scores. The results of labeling data were divided into 2 classes, namely the positive class with 1676 and the negative class with 5324. Classification and evaluation were carried out using the *Naïve Bayes* method. Based on the comparison value of training and testing data of 8:2, the results obtained were an accuracy level of 84%, then precision of 87% and recall of 92%.

Keywords—*Netflix*; Sentiment Analysis; *Google Playstore*; *Machine Learning*; *Naïve Bayes*



I. PENDAHULUAN

Perkembangan teknologi informasi sangatlah cepat yang didorong oleh kebutuhan akan data dan informasi berbasis komputer sehingga aplikasi *streaming* berkembang dengan sangatlah pesat (Nur Alamsyah 2020). Kegiatan menonton film melalui aplikasi *streaming online* masih banyak disukai oleh masyarakat saat ini hingga menjadi kebiasaan yang dilakukan ketika masyarakat merasa jenuh bahkan *streaming online* (Indriani et al. 2023). *Netflix* adalah layanan *streaming* berbasis langganan yang memungkinkan para pengguna dapat menonton acara TV dan film tanpa iklan di perangkat yang terhubung ke internet menggunakan perangkat seperti (*Smartphone, smartTV, tablet, PC dan laptop*) kapan pun dan dimanapun. Didirikan pada tanggal 1997 di *Scotts Valley, California* dan masuk di Indonesia pada tahun 2016, *Netflix* menyediakan arsip atau kumpulan film paling lengkap dengan wilayah pengoperasian terbanyak di dunia (Fani Al-shufi and Erfina 2021a).

Data yang dihasilkan dari aplikasi *netflix* bukan hanya terkait tentang data perfilman, melainkan ada data lain yang dihasilkan dari proses distribusi aplikasi, yaitu *Google Play Store*. Setiap pengguna yang mengunduh aplikasi *netflix* dapat memberikan ulasan terhadap aplikasi melalui fitur yang sudah disediakan oleh *Website Google Play Store*. Ulasan tersebut dapat diakses secara bebas oleh siapa saja data ulasan dari aplikasi *netflix* dapat bermanfaat apabila diolah dengan baik (Anon n.d.-d). Hasil pengolahan data akan menghasilkan masukan untuk perbaikan dan pengembangan aplikasi, karena kritik dan saran terbaik berasal dari pendapat langsung semua pengguna. Pendapat dari setiap pengguna pada

Google Play Store dapat berpengaruh pada calon pengguna sebagai bahan pertimbangan penggunaan suatu aplikasi. Namun, dengan banyaknya data ulasan yang ada, akan sulit diolah secara manual serta sangat kesulitan untuk mendapatkan hasilnya. Analisis sentimen ialah proses mengekstraksi, mengolah dan memahami data berupa teks yang tidak terstruktur secara otomatis guna mengambil informasi sentimen yang terdapat pada sebuah kalimat pendapat atau opini. Penilaian yang dilakukan terhadap data teks dari berbagai sudut pandang untuk mengetahui beberapa opini positif atau negatif, dari teks yang banyak dengan konteks yang beragam (Purba and Rizki Padya 2023).

Berdasarkan latar belakang masalah tersebut maka dilakukan penelitian untuk mengetahui opini masyarakat tentang aplikasi *streaming* film di *Play Store* (Djamzuri and Putra Mulyana 2022). Hal ini dapat bermanfaat untuk meningkatkan kualitas layanan dapat dengan memanfaatkan *review* atau ulasan dari pengguna aplikasi *Netflix* yang terdapat pada di *Google Play Store*, di dalam ulasan yang di berikan terdapat banyak pengalaman pengguna aplikasi yang baik maupun yang buruk, positif maupun negatif, atau sentimen penggunaannya, selain itu ulasan dari pengguna juga dapat memberikan pengaruh bagi calon pengguna. (Sudarsono et al. 2021)

Penelitian ini akan menjabarkan proses yang akan menghasilkan analisis sentimen dengan menerapkan algoritma *Naive Bayes* dengan *text mining rapid miner*. Algoritma *Naive Bayes* akan dimanfaatkan sebagai metode untuk menghasilkan prediksi suatu ulasan, apakah berada dalam Kategori positif atau negatif. Algoritma *Naive Bayes* dipilih karena telah banyak digunakan pada proses *text mining* dan akurasi tinggi. Dengan

penerapan algoritma *Naive Bayes* juga dapat mempersingkat waktu pengklasifikasian dalam analisis sentimen.(Fani Al-shufi and Erfina 2021b)

Dari uraian serta penjelasan yang telah dijelaskan diatas, maka penulis tertarik untuk melakukan penelitian dengan judul “**Analisis Sentimen pengguna Aplikasi Netflix pada Google Playstore menggunakan metode Naïve Bayes**”.

II. TINJAUAN PUSTAKA

A. Pengertian Analisis

Analisis adalah penyelidikan terhadap suatu peristiwa (karangan, perbuatan, dsb) untuk mengetahui keadaan yang sebenarnya (sebab, musabab, duduk perkaranya, dsb) penguraian suatu pokok atau berbagai bagiannya dan penelaahan bagian itu sendiri serta hubungan antar bagian untuk memperoleh pengertian yang tepat dan pemahaman arti keseluruhan.(Baiti and Iswandi 2022)

B. Pengertian Analisis Sentimen

Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negatif. Analisis sentimen diberikan oleh pengguna internet pada media sosial untuk memberikan suatu penilaian atau opini pribadi.[10]

C. Aplikasi

Aplikasi adalah istilah yang digunakan untuk pengguna komputernya bagi pemecahan masalah. Biasanya istilah aplikasi dipasangkan atau digabungkan dengan suatu perangkat lunak. Aplikasi berasal dari kata *aplication* yang artinya penerapan, lamaran, penggunaan.

Secara istilah aplikasi adalah program siap pakai yang direka untuk melaksanakan suatu fungsi bagi pengguna atau aplikasi yang lain digunakan untuk satu tujuan. (Anon n.d.-b)

D. Netflix

Netflix sebagai salah satu media layanan streaming video film dan serial televisi dunia memberikan dampak kepada budaya perfilman umat manusia. Penggunaan *netflix* menjadi sebuah *new habits/kebiasaan* baru dalam menikmati film-film terbaru baik lokal maupun film import. Masifnya gawai telepon pintar menjadikan *streaming* video on demand (SVoD) sebagai *alternative* baru pengalaman *Google Play Store*.(Tanggraeni and Sitokdana 2022)



Gambar 1. Netflix

E. Google PlayStore

Saat ini telah tersedia sekitar 1.200.000 aplikasi *mobile* pada *Gogle Play*. Dengan sejumlah aplikasi tersebut membuat pengguna memiliki banyak pilihan. Untuk mengatasi masalah tersebut, pertama-tama mereka harus mengetahui fungsi dari aplikasi yang akan di-*install*, bagaimana pendapat konsumen lain terhadap aplikasi tersebut dan kemudian memutuskan membeli atau tidak. Biasanya, ulasan pengguna berisi dua bagian, yaitu nilai rating dan komentar secara tekstual. Nilai



rating menunjukkan evaluasi keseluruhan pengalaman pengguna menggunakan skala numerik, namun komentar tekstual mampu bercerita lebih mendalam. (Gormantara 2020)

F. *Naïve Bayes*

Naïve Bayes Classifier (NBC) *Naïve Bayes* merupakan sebuah metode pengklasifikasian dengan menggunakan probabilitas sederhana yang berakar pada Teorema *Bayes* dan memiliki asumsi ketidaktergantungan (*independent*) yang tinggi dari masing – masing kondisi atau kejadian. (Iqbal Ahmadi et al. n.d.)

G. *Text Mining*

Text mining adalah suatu langkah dari analisis teks yang secara otomatis dilakukan oleh komputer dengan tujuan menggali informasi yang berkualitas dari suatu rangkaian teks yang terangkum dalam sebuah dokumen. *Text Mining* merupakan salah satu proses yang bidang dipakai untuk mengerjakan 20 klasifikasi dokumen yang berusaha untuk menemukan bentuk yang menarik dari sekumpulan data yang terkumpul. (Fakhri Irfani and Triyanto n.d.)

H. *Preprocessing*

Pre Processing merupakan tahapan awal dalam melakukan pengolahan teks yang digunakan untuk mengubah bentuk dokumen menjadi data terstruktur sesuai kebutuhannya agar dapat diolah lebih lanjut dalam proses *text mining*. Tahapan *pre process* dalam klasifikasi bertujuan untuk meningkatkan akurasi klasifikasi data. (Laia and Yamin 2023)

I. *Klasifikasi*

Klasifikasi merupakan model yang memprediksi suatu nilai yang tidak diketahui. Model klasifikasi terbagi menjadi dua yaitu, melatih data kemudian membedakan kedalam *class* dan melakukan evaluasi terhadap objek

tidak dikenal. Klasifikasi memerlukan persiapan data yang baik dan pemilihan bahasa pemrograman yang mampu mendukung *machine learning*. Salah satu tugas yang dapat dilakukan dengan data mining adalah pengklasifikasian. Dalam klasifikasi terdapat target variabel kategori. Metode – metode yang telah dikembangkan oleh periset untuk menyelesaikan kasus klasifikasi antara lain sebagai berikut ini : (Adi Cakranegara et al. 2022)

1. Pohon Keputusan
2. *Naïve Bayes*
3. Jaringan Saraf Tiruan
4. Analisis Statistik
5. Algoritma Genetik
6. *Rough Sets*
7. Pengklasifikasi *K-nearest Neighbour*
8. Metode Berbasis Aturan
9. *Memory Based Reasoning*
10. *Support Vector Machine*

J. *Python*

Python dikenal sebagai bahasa pemrograman *interpreter*, karena *Python* dieksekusi dengan sebuah interpreter. *Python* saat ini telah menduduki posisi 28 4 atau 5 bahasa pemrograman yang paling sering digunakan di seluruh dunia. Bahasa pemrograman *Python* memiliki beberapa fitur yang dapat digunakan oleh pengembang perangkat lunak. Berikut adalah beberapa fitur yang ada pada bahasa pemrograman *Python*:

1. *Multi Paradigm Design*
2. *Open Source*
3. *Simplicity*
4. *Library Support*
5. *Portability*
6. *Extendable*
7. *Scalability*

Dalam melakukan analisis data, *Python*



memiliki *library* yang sering digunakan yaitu:

1. NumPy merupakan *library* yang digunakan untuk melakukan perhitungan matematika maupun perhitungan ilmiah.
2. Matplotlib merupakan *library* yang digunakan untuk visualisasi data
3. Pandas sering digunakan untuk memindahkan data saat melakukan analisis maupun *Remove Punctuation* data dan melakukan rangkuman statistic
4. Scraper atau yang lebih dikenal Scraping adalah *tools* atau *library* pada Python yang digunakan untuk mengumpulkan data dari *Goggle Play Store*.

III. METODOLOGI PENELITIAN

a. Metode Penelitian

Metode penelitian yang digunakan adalah metode analisis deskriptif dengan pendekatan kuantitatif, kuantitatif adalah suatu proses menemukan pengetahuan yang menggunakan data berupa angka sebagai alat menganalisis keterangan mengenai apa yang ingin kita ketahu (Anon n.d.-a).

Penelitian ini dilakukan pada *Google Play*, sedangkan untuk data yang diambil yaitu mulai pada maret 2023 sampai dengan desember 2023. Pada penelitian ini penulis menggunakan beberapa metode untuk mendapatkan data dan informasi untuk menyelesaikan permasalahan yang ada.

b. Metode Pengumpulan Data

Pengumpulan data dilakukan untuk mendapatkan suatu informasi yang dibutuhkan dalam mencapai tujuan dari penelitian ini. Dalam penelitian ini penulis menggunakan beberapa metode, adapun metodenya yaitu sebagai berikut :

1. Studi pustaka merupakan metode pengumpulan data kepastakaan yang dilakukan dengan mengumpulkan melalui jurnal, *literature*, buku, maupun berbagai situs internet yang kemudian dijadikan sebagai sumber pustaka yang berkaitan dengan materi penelitian ini terutama tentang analisi sentimen menggunakan Algoritma *Naive Bayes* (Athallah Muhammad et al. n.d.).
2. Data Primer Merupakan Pengumpulan data pada penelitian ini menggunakan salah satu *library* yang ada pada *pyhton* yaitu *scraper*. *Scraper* bisa mengambil data di aplikasi *Netflix* secara langsung. data yang diperoleh untuk penelitian ini yaitu data tentang ulasan *rating*, ulasan komentar terhadap salah satu aplikasi *Netflix*.

c. Metode Analisa

Metode analisa yang digunakan dalam penelitian ini adalah metode Algoritma *Naive Bayes*. *Teorema Bayes* merupakan dasar dari teknik pengklasifikasian yang diterapkan pada Klasifikasi *Naive Bayes*. *Teorema Bayes* sendiri ditemukan oleh ilmuwan inggris bernama *Thomas Bayes*. dalam melakukan analisis sentimen banyak yang menerapkan Algoritma *Naive Bayes* sebagai metodenya. Algoritma *Naive Bayes* memiliki keunggulan, yaitu efisien karna mampu membuat proses analisis sentimen menjadi singkat. Selain itu, Algoritma *Naive Bayes* memiliki akurasi yang cenderung tinggi walaupun memiliki data Jatih yang sedikit. Pada penelitian penulis menggunakan satu nilai perbandingan dari data *training* dan data *testing* yaitu: Data *Training* 80% dan *testing* 20% (8:2). Penerapan metode *Naive Bayes* dilakukan dengan cara Semanggilan *library*



Naive Bayes pada python. Penerapan *Naive Baves* dibarengi dengan pengujian menggunakan *confusion matrix* yang menghasilkan *matrix* yang berisi ulasan *true positif* (benar positif), *False positif* (salah positif), *true negatif* (benar negatif), dan *false negatif* (salah negatif) untuk mencari nilai akurasi, presisi dan *recall*. (Daryfayi, Daulay, and Asror n.d.)

d. *Metode Pengujian Dan Pengolahan Data*

1. *Metode Pengujian*

Dalam penelitian ini metode pengujian dilakukan menggunakan *library* yang terdapat pada metode Algoritma *Naïve Bayes* yaitu *confusion matrix*. Pada dasarnya pengujian menggunakan *confusion matrix* 4 (empat) istilah yang

1. Akurasi adalah tingkat kedekatan nilai sebenarnya yang didapatkan. Berikut adalah rumus untuk menentukan akurasi:

$$\text{Akurasi} = \frac{TN + TP}{TN + FN + TP + FP}$$

Keterangan:

TN = *True Negative*

TP = *True Positive*

FN = *False Negative*

TN = *True Negative*

2. Presisi merupakan sebuah tingkat kecocokan dari data yang diambil dengan informasi yang ada. Berikut adalah rumus untuk menentukan presisi:

$$\text{Presisi} = \frac{TP}{TP + FP}$$

3. *Recall* adalah tingkat kedekatan dari nilai sebenarnya dengan nilai yang didapatkan. Berikut rumus untuk menghitung *Recall*:

$$\text{Recall (N)} = \frac{TP}{TN + FN}$$

$$\text{Recall (P)} = \frac{TP}{TN + FP}$$

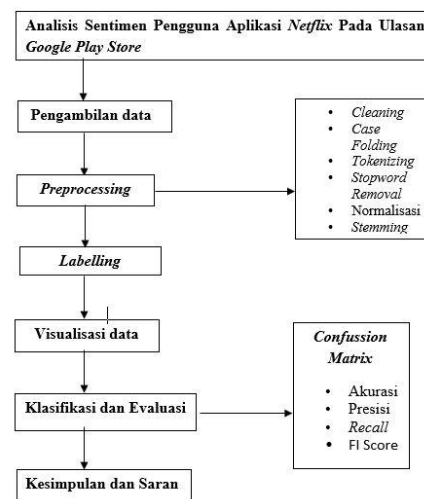
4. *F1 Score* merupakan suatu perbandingan rata-rata presisi dan *recall* yang telah dibobotkan. Berikut rumus untuk menghitung *F1 Score*:

$$\text{F-1 Score} = 2X$$

$$\frac{\text{Recall} \times \text{Presisi}}{\text{Recall} + \text{Presisi}}$$

e. *Pengolahan Data*

Pada penelitian ini data yang berhasil didapatkan dari *Google Play* pada aplikasi *Netflix* akan dilakukan *labelling* dan *preprocessing*. Adapun tahapan dalam *labelling* dan *preprocessing* adalah sebagai berikut :



Gambar 2. Proses *Preprocessing*

1. *Cleaning*

Cleaning adalah proses



penghilangan komponen tertentu yang terdapat dalam tweet yakni *Uniform Resource Locator* (URL), *username*, RT (*Retweet*), karakter HTML, dan hashtag (Anon n.d.-c).

2. *Case folding*

Case Folding adalah proses penyamaan case dalam sebuah dokumen. Hal ini dilakukan untuk mempermudah pencarian. Tidak semua dokumen teks konsisten dalam penggunaan huruf kapital. Oleh karena itu peran *case folding* dibutuhkan dalam mengkonversi keseluruhan teks dalam dokumen menjadi suatu bentuk standar (dalam hal ini huruf kecil atau *lowercase*).

3. *Stopword Removal*

Stopword Removal merupakan proses penyaringan atau penghapusan kata yang sering muncul, namun tidak mengubah makna dari suatu ulasan dari tokenisasi.

4. *Tokenizing*

Tokenizing adalah tahap penguraian teks yang semula berupa kalimat-kalimat yang berisi kata-kata. proses pemotongan *string* berdasarkan tiap kata yang menyusunnya, umumnya setiap kata akan terpisahkan dengan karakter spasi..

5. Normalisasi

Pada tahap ini merupakan tahap untuk mengubah kalimat yang tidak baku atau *slangword* menjadi kalimat baku yang sesuai KBBI (Kamus Besar Bahasa Indonesia).

6. *Stemming*

Terakhir adalah tahap *stemming*. *Stemming* merupakan tahap yang dilalui untuk membuat kata yang memiliki imbuhan menjadi kata dasar. Kata imbuhan berupa awalan, sisipan, maupun akhiran akan dihilangkan melalui tahap ini. Untuk ulasan yang menggunakan bahasa indonesia, *stemming* dilakukan dengan menggunakan *library* Sastrawi pada python (Wahyudi et al. 2021).

IV. HASIL DAN PEMBAHASAN

Pengumpulan data dilakukan menggunakan *library snsrape* yang ada pada bahasa pemrograman *python*. Langkah pertama dalam pengumpulan data adalah menginstal *library snsrape* di *google colab*, dengan menggunakan *source code* berikut ini “!pip install -q snsrape”. Kemudian dapat dilakukan pengambilan data sesuai dengan kata kunci dan jangka waktu yang penulis inginkan. Berikut adalah gambar *source code* yang penulis gunakan untuk melakukan pengambilan data.

Untuk mendapatkan hasil dalam penelitian ini diperlukan beberapa tahapan yang dilakukan seperti tahap pengumpulan data dengan *Scrapping* data ulasan, tahap *preprocessing*, tahap *labelling*, klasifikasi dan evaluasi.

Tahap Pengumpulan Data

Pada penelitian ini data yang dikumpulkan berjumlah 7000 data ulasan *google playstore* pada aplikasi *netflix*. Pengumpulan data dilakukan menggunakan kata kunci yaitu “ulasan pengguna aplikasi *netflix*” dan jangka waktu pengumpulan data adalah dari maret 2023 sampai dengan desember 2023. Data yang dikumpulkan terdapat dua kolom yaitu *index*, *username*,



score, at dan content, kemudian disimpan dengan format csv.

Tahap Preprocessing

Tahap *preprocessing* adalah mengubah data yang berhasil dikumpulkan menjadi format yang diperlukan. Tahap *preprocessing* dilakukan dengan bahasa pemrograman *python* yang dijalankan di *goolge colab*. Berikut merupakan tahapan³. *Preprocessing*:

1. Cleaning

Cleaning yaitu proses pembersihan data ulasan *google playstore* pada Aplikasi *Netflix* yang tidak dibutuhkan seperti karakter HTML, *username* (@username), *emoticon*, *hastag*(#), angka, url (<https://tommyganteng.com>) dan lain-lain. Adapun hasil dari proses *cleaning* dari data ulasan *google playstore* pada Aplikasi *Netflix* yang sudah ada didapatkan dapat dilihat pada gambar berikut ini:

username	score	at	content	text_casefolding
0	amisa Ulla	2024-01-03 10:28:58	Pas mutar video jd slow mo. Suara sama gambar.	pas mutar video jd slow mo suara sama gambar g...
1	Alifah Hasna	2024-01-03 07:49:15	Jelek bgt aplikasinya baru buka malah disuruh...	jelek bgt aplikasinya baru buka malah disuruh...
2	Iqbal hormansyah	2024-01-03 07:00:04	Ini gimana yah saya langganan pake paket stand...	ini gimana yah saya langganan pake paket stand...
3	fadurrahman akbar	2024-01-02 13:17:50	Janganakan di putar di tv, di putar di hp ga bi...	janngankan di putar di tv di putar di hp ga bis...
4	akatsuno puji	2024-01-02 12:33:54	lebih disaring lagi filmnya ada yang tidak pantas	lebih disaring lagi filmnya ada yang tidak pantas
5	luki 90	2024-01-02 11:32:34	Frame drop kalau lagi nonton, baik streaming m...	frame drop kalau lagi nonton baik streaming ma...
6	Raynaldo Carsella	2024-01-02 10:47:09	Infinitx zero30 saat play video langsung patah...	infinitx zero saat play video langsung patah ga...
7	Azri Maulidiya	2024-01-02 07:45:14	Ribet bgt loginyanya, udi di taukan malah ngula...	ribet bgt loginyanya udi di taukan malah ngulan...
8	Cecap Rukmana	2024-01-02 07:37:11	Konten-kontennya saya suka, cuman yang mengganja...	kontenkontennya saya suka cuman yang mengganja...

Gambar 3. Hasil Cleaning

2. *Case Folding* merupakan proses yang digunakan untuk menyeragamkan huruf yang terdapat pada ulasan *google playstore* pada Aplikasi *Netflix*, dari huruf kapital menjadi huruf kecil (*Lower Case*). Adapun hasil dari proses *casefolding* dari data ulasan *google playstore* pada Aplikasi *Netflix* yang sudah ada didapatkan dapat dilihat pada gambar berikut ini:

username	score	at	content	text_casefolding	text_clean	text_stopword
0	amisa Ulla	2024-01-03 10:28:58	Pas mutar video jd slow mo. Suara sama gambar.	pas mutar video jd slow mo suara sama gambar g...	pas mutar video jd slow mo suara sama gambar g...	pas mutar video jd slow mo suara sama gambar g...
1	Alifah Hasna	2024-01-03 07:49:15	Jelek bgt aplikasinya baru buka malah disuruh...	jelek bgt aplikasinya baru buka malah disuruh...	jelek bgt aplikasinya baru buka malah disuruh...	jelek bgt aplikasinya baru buka malah disuruh...
2	Iqbal hormansyah	2024-01-03 07:00:04	Ini gimana yah saya langganan pake paket stand...	ini gimana yah saya langganan pake paket stand...	ini gimana yah saya langganan pake paket stand...	gimana yah langganan pake paket stand...
3	fadurrahman akbar	2024-01-02 13:17:50	Janganakan di putar di tv, di putar di hp ga bi...	janngankan di putar di tv di putar di hp ga bis...	janngankan di putar di tv di putar di hp ga bis...	putar tv putar hp ga udi mngga pas disang...
4	akatsuno puji	2024-01-02 12:33:54	lebih disaring lagi filmnya ada yang tidak pantas	lebih disaring lagi filmnya ada yang tidak pantas	lebih disaring lagi filmnya ada yang tidak pantas	disaring filmnya
5	luki 90	2024-01-02 11:32:34	Frame drop kalau lagi nonton, baik streaming m...	frame drop kalau lagi nonton baik streaming ma...	frame drop kalau lagi nonton baik streaming ma...	frame drop kalau lagi nonton baik streaming ma...
6	Raynaldo Carsella	2024-01-02 10:47:09	Infinitx zero30 saat play video langsung patah...	infinitx zero saat play video langsung patah ga...	infinitx zero saat play video langsung patah ga...	infinitx zero saat play video langsung patah ga...
7	Azri Maulidiya	2024-01-02 07:45:14	Ribet bgt loginyanya, udi di taukan malah ngula...	ribet bgt loginyanya udi di taukan malah ngulan...	ribet bgt loginyanya udi di taukan malah ngulan...	ribet bgt loginyanya udi di taukan malah ngulan...
8	Cecap Rukmana	2024-01-02 07:37:11	Konten-kontennya saya suka, cuman yang mengganja...	kontenkontennya saya suka cuman yang mengganja...	kontenkontennya saya suka cuman yang mengganja...	kontenkontennya saya suka cuman yang mengganja...

Gamabr 4. Hasil Case Folding

Stopword Removal adalah proses penyaringan dan penghapusan kata yang tidak penting dari *tokenizing*. Pada penelitian ini penulis menggunakan *stopword* bahasa indonesia yang didapatkan dari *library* NLTK yang dapat pada bahasa pemograman *Python* kemudian ditambah dengan *Stopword* yang penulis buat sendiri. Adapun hasil dari proses *stopword removal* dari data ulasan *google playstore* pada Aplikasi *Netflix* yang sudah ada didapatkan dapat dilihat pada gambar berikut ini:

username	score	at	content	text_casefolding	text_clean	text_stopword
0	amisa Ulla	2024-01-03 10:28:58	Pas mutar video jd slow mo. Suara sama gambar.	pas mutar video jd slow mo suara sama gambar g...	pas mutar video jd slow mo suara sama gambar g...	pas mutar video jd slow mo suara sama gambar g...
1	Alifah Hasna	2024-01-03 07:49:15	Jelek bgt aplikasinya baru buka malah disuruh...	jelek bgt aplikasinya baru buka malah disuruh...	jelek bgt aplikasinya baru buka malah disuruh...	jelek bgt aplikasinya baru buka malah disuruh...
2	Iqbal hormansyah	2024-01-03 07:00:04	Ini gimana yah saya langganan pake paket stand...	ini gimana yah saya langganan pake paket stand...	ini gimana yah saya langganan pake paket stand...	gimana yah langganan pake paket stand...
3	fadurrahman akbar	2024-01-02 13:17:50	Janganakan di putar di tv, di putar di hp ga bi...	janngankan di putar di tv di putar di hp ga bis...	janngankan di putar di tv di putar di hp ga bis...	putar tv putar hp ga udi mngga pas disang...
4	akatsuno puji	2024-01-02 12:33:54	lebih disaring lagi filmnya ada yang tidak pantas	lebih disaring lagi filmnya ada yang tidak pantas	lebih disaring lagi filmnya ada yang tidak pantas	disaring filmnya

Gambar 5. Hasil Stopword Removal

- d) *Tokenzing* merupakan tahapan dimana proses ini dilakukan untuk memisahkan kalimat untuk menjadi kata-perkata. Adapun hasil dari proses *Tokenzing* dari data ulasan *google playstore* pada Aplikasi *Netflix* yang sudah ada didapatkan dapat dilihat pada gambar berikut ini:

username	score	at	content	text_casefolding	text_clean	text_stopword	text_tokens
0	amisa Ulla	2024-01-03 10:28:58	Pas mutar video jd slow mo. Suara sama gambar.	pas mutar video jd slow mo suara sama gambar g...	pas mutar video jd slow mo suara sama gambar g...	pas mutar video jd slow mo suara sama gambar g...	[pas, mutar, video, jd, slow, mo, suara, sama, gambar, g...]
1	Alifah Hasna	2024-01-03 07:49:15	Jelek bgt aplikasinya baru buka malah disuruh...	jelek bgt aplikasinya baru buka malah disuruh...	jelek bgt aplikasinya baru buka malah disuruh...	jelek bgt aplikasinya baru buka malah disuruh...	[jelek, bgt, aplikasinya, baru, buka, malah, disuruh, ...]
2	Iqbal hormansyah	2024-01-03 07:00:04	Ini gimana yah saya langganan pake paket stand...	ini gimana yah saya langganan pake paket stand...	ini gimana yah saya langganan pake paket stand...	gimana yah langganan pake paket stand...	[gimana, yah, langganan, pake, paket, stand, ...]
3	fadurrahman akbar	2024-01-02 13:17:50	Janganakan di putar di tv, di putar di hp ga bi...	janngankan di putar di tv di putar di hp ga bis...	janngankan di putar di tv di putar di hp ga bis...	putar tv putar hp ga udi mngga pas disang...	[putar, tv, putar, hp, ga, udi, mngga, pas, d...]
4	akatsuno puji	2024-01-02 12:33:54	lebih disaring lagi filmnya ada yang tidak pantas	lebih disaring lagi filmnya ada yang tidak pantas	lebih disaring lagi filmnya ada yang tidak pantas	disaring filmnya	[disaring, filmnya]

Gambar 6. Hasil Tokenzing



- e) Normalisasi Pada tahap ini merupakan tahap untuk mengubah kalimat yang tidak baku atau *slangword* menjadi kalimat baku. Adapun hasil dari proses normalisasi dari data ulasan yang sudah didapatkan dapat dilihat pada gambar berikut ini

userName	score	at	content	text_clean	text_standard	text_tokens	normalisasi
0	Ravenna Fabian ananta kati	5	2024-01-08 23:08:51	Tolong dong diperbaiki mengapa masuk harus pun...	tolong dong diperbaiki mengapa masuk harus pun...	tolong dong diperbaiki masuk, harus kreditoranpe...	tolong diperbaiki masuk, harus kreditoranpe...
1	07x	1	2024-01-08 12:31:58	lebih baik dan terakut juga uninstal apk ini...	lebih baik dan terakut juga uninstal apk ini...	uninstal apk ini sampai tuntas baw...	Perbaiki, uninstal, apk, ini, sampai, bereslata...
2	Nanda Ekhia	2	2024-01-08 11:23:28	Apaan anji di update bukan nya bagus malah jelek	Apaan anji di update bukan nya bagus malah jelek	anjir update nya bagus jelek	[anjir, update, nya, bagus, jelek]
3	Nabela Yansary	2	2024-01-08 04:21:22	Ini kenapa alun saya paksa diinstal udah di do	Ini kenapa alun saya paksa diinstal udah di do	alun paksa putar udah download gk paksa put...	[alun, paksa, diinstal, udah, download, gk, g...
4	Mula Ahyakaja	5	2024-01-08 01:13:49	Bagus hanya banyak film yg ter di sala kir bi...	bagus hanya banyak film yg ter di sala kir bi...	bagus film yg out of date kir bahwa film	[bagus, film, yg, out, of, date, kir, bahwa, film]

Gambar 7. Hasil Normalisasi

- f) *stemming* adalah proses untuk menghilangkan awalan atau akhiran kata yang terdapat kata sambung, kata depan, kata ganti, menjadi kata dasar. Proses *Stemming* dilakukan dengan *library* sastrawi yang terdapat pada bahasa pemrograman *python*. Adapun hasil dari proses *Stemming* dari data ulasan *google playstore* pada Aplikasi *Netflix* yang sudah ada didapatkan dapat dilihat pada gambar berikut ini:

userName	score	at	content	text_cleanfolding	text_clean	text_standard	text_tokens
0	ananta Ulla	3	2024-01-03 10:26:50	Pas mutar video jd slow mo. Suara sama gambar g...	pas mutar video jd slow mo mo suara sama gambar g...	pas mutar video jd slow mo suara sama gambar g...	[pas, mutar, video, jd slow, mo, suara, gambar, g...
1	Alifah Hima	2	2024-01-03 07:49:15	Jelek lagi aplikasinya baru buka malah disuruh...	jelek lagi aplikasinya baru buka malah disuruh...	jelek lagi aplikasinya baru buka malah disuruh...	[jelek, lagi, aplikasinya, baru, buka, malah, disuruh, ...]
2	laila hormanyah	3	2024-01-03 07:09:04	Ini gimana yeh saya langganan pake paket stand...	ini gimana yeh saya langganan pake paket stand...	gimana yeh langgan pake paket standart engga s...	[gimana, yeh, langgan, pake, paket, standart, engga, s...
3	Indarrahman akbar	1	2024-01-02 13:17:50	Jangan ker di putar di tv di putar di hp ga bi...	jeangan ker di putar di tv di putar di hp ga bi...	putar tv putar hp ga udh minggu pas doang	[putar, tv, putar, hp, ga, udh, minggu, pas, doang]
4	aklatara paji	1	2024-01-02 12:23:54	lebi diasing lagi filmnya ada yang tidak pantas	lebi diasing lagi filmnya ada yang tidak pantas	asing film	[asing, film]

Gambar 8 Hasil Dari Stemming

- g) Setelah tahap preprocessing selesai, maka tahap selanjutnya adalah *labelling*. Proses *Labelling* dilakukan membagi ke dalam dua kelas sentimen berdasarkan *score rating* ulasan pada *google playstore*, yaitu jika data mendapat skor 1,2 atau 3 maka termasuk kelas negatif, dan jika mendapat skor 4 dan 5, maka masuk pada kelas positif. Adapun hasil dari proses *Labelling*

dari data ulasan *google playstore* pada Aplikasi *Netflix* yang sudah ada didapatkan dapat dilihat pada gambar berikut ini:

text_tokens	text_standard	label
[pas, mutar, video, jd, slow, mo, suara, gamba...]	pas mutar video jd slow mo suara gambar gak ba...	negatif
[jelek, bgt, aplikasinya, buka, disuruh, bayar...]	jelek bgt aplikasi buka suruh bayar trus ngele...	negatif
[gimana, yah, langganan, pake, paket, standart...]	gimana yah langgan pake paket standart engga s...	negatif
[putar, tv, putar, hp, ga, udh, minggu, pas, d...]	putar tv putar hp ga udh minggu pas doang	negatif
[disaring, filmnya]	saring film	negatif

Gambar 9. Hasil Labelling

- h) Setelah dilakukan labelling data, proses selanjutnya yaitu klasifikasi dan evaluasi. Data yang sudah diberi label kemudian dibagi menjadi dua, yaitu data *training* dan data *testing*. Proses klasifikasi dan evaluasi pada penelitian ini dilakukan menggunakan metode *Naïve Bayes*.

- i) Pembagian Data *Training* dan Data *Testing*

Langkah awal klasifikasi menggunakan metode *Naïve Bayes* adalah membagi data menjadi dua yaitu data *training* dan data *testing*. Pada penelitian ini penulis menggunakan perbandingan 80% : 20%.

- j) Klasifikasi dan Evaluasi menggunakan *Naïve Bayes*

Selanjutnya klasifikasi dan evaluasi menggunakan metode Algoritma *Naïve*



Bayes, dimana pada metode ini akan dilakukan pelatihan dengan cara mendeteksi ciri-ciri data pada tiap kelas dari data *training*. Dan dari hasil pelatihan pada metode Algoritma *Naïve Bayes* yang kemudian akan dievaluasi menggunakan data testing. proses evaluasi dilakukan menggunakan *library* yang terdapat pada metode Algoritma *Naïve Bayes* yaitu *confusion matrix* untuk mencari nilai dari akurasi, presisi, dan *recall*. Berikut adalah hasil dari proses klasifikasi menggunakan metode Algoritma *Naïve Bayes* yang dapat dilihat pada gambar :

	precision	recall	f1-score	supp
negatif	0.88	0.91	0.89	1
positif	0.68	0.60	0.64	1
accuracy			0.84	1
macro avg	0.78	0.76	0.77	1
weighted avg	0.83	0.84	0.83	1

Gambar 10. Hasil Klasifikasi Metode *Naïve Bayes*

k) Pembahasan

Pada penelitian ini penulis membahas masalah tentang Analisis Sentimen Pengguna Aplikasi *Netflix* Pada Ulasan *Google Playstore* Menggunakan Metode *Naïve Bayes*. Langkah pertama yaitu dimulai dengan pengambilan data ulasan pada website *google play* yaitu aplikasi *Netflix*. Pengambilan data dilakukan menggunakan *library scraper*, dan untuk data ulasan aplikasi *Netflix* akan diambil mulai dari Maret 2023 sampai dengan Desember 2023. Ulasan dari pengguna aplikasi *Netflix* yang berhasil didapat berjumlah 7000 ulasan. Kemudian proses selanjutnya yaitu *labelling* dimana ulasan yang berhasil dikumpulkan

akan diberi label positif dan negatif. Untuk pelebelannya sendiri yaitu 3 akan berlabel negatif dan untuk *ratting* 3 itu adalah positif. Langkah selanjutnya yaitu *preprocessing* data, setelah dilakukan pengambilan data dan *labelling* data, Data ulasan yang berhasil didapatkan berjumlah 7000 data, yang kemudian dilakukan *preprocessing*. Pada tahap *preprocessing* dari 7000 data yang berhasil didapatkan menghasilkan 6999 data yang berhasil lolos dari tahap *preprocessing*.

Proses selanjutnya yaitu *labelling* data, *labelling* data dilakukan berdasarkan *score rating* data ulasan *google play store*. Hasil dari *labelling* data yaitu 5324 untuk sentimen negatif dan 1676 untuk sentimen positif. Setelah dilakukan *labelling* data maka proses selanjutnya yaitu klasifikasi dan evaluasi menggunakan metode *Naïve Bayes*. Klasifikasi dan evaluasi menggunakan metode *Naïve Bayes* diawali dengan membagi data menjadi data *training* dan data *testing*. Pada penelitian ini perbandingan data *training* dan data *testing* yang digunakan yaitu 8:2 (80% : 20%).

Pembagian data dilakukan dengan data *training* 80% (5600 data) dan data *testing* 20% (1400 data). Kemudian menerapkan metode *Naïve Bayes* pada pembagian data tersebut sehingga memiliki hasil akurasi sebesar 84% dengan nilai presisi 88% beserta *recall* 91%

V. KESIMPULAN

Adapun kesimpulan dari penelitian yang telah dilakukan mengenai analisis sentimen aplikasi youtube di *google playstore* menggunakan *Naïve Bayes* adalah sebagai berikut :



1. Data yang digunakan dalam penelitian ini diperoleh dengan memanfaatkan *library snsrape*, sebanyak 7000 data dan dilakukan *preprocessing* sehingga tersisa menjadi 6999 data yang bisa diolah. Data yang diambil merupakan data ulasan di *google playstore* dari jangka waktu bulan maret 2023 sampai desember 2023.
2. Hasil *labelling* 2 kelas data yaitu kelas positif sebanyak 1676, dan kelas negatif sebanyak 5324.
3. Kesimpulan proses akhir klasifikasi dan penerapan algoritma *Naïve Bayes* menggunakan perbandingan data *training* dan data *testing* yang telah penulis lakukan berdasarkan data *training* 80% dan data *testing* 20% menghasilkan nilai akurasi yaitu 84% dengan nilai presisi 88% beserta *recall* 91%.

VI. DAFTAR REFERENSI

- Adi Cakranegara, Pandu, Yanti Budiasih, Hendy Tannady, and Yohanes Totok Suyoto. 2022. *Analysis Of The Effect Of Perceived Value And Brand Image On Netflix Service Purchase Decisions Analisis Pengaruh Perceived Value Dan Brand Image Terhadap Purchase Decision Layanan Netflix*. Vol. 3.
- Anon. n.d.-a. '224-Article Text-794-1-10-20191010'.
- Anon. n.d.-b. *ANALISIS SENTIMEN PADA TWITTER MENGGUNAKAN METODE MAXIMUM ENTROPY DAN SUPPORT VECTOR MACHINE*.
- Anon. n.d.-c. *ANALISIS SENTIMEN PADA TWITTER MENGGUNAKAN METODE MAXIMUM ENTROPY DAN SUPPORT VECTOR MACHINE*.
- Anon. n.d.-d. 'WHATSAPP' MEDIA KOMUNIKASI EFEKTIF MASA KINI (STUDI KASUS PADA LAYANAN JASA INFORMASI ILMIAH DI KAWASAN PUSPIPTEK) Rahartri Pusat Data Dan Dokumentasi Ilmiah-Lembaga Ilmu Pengetahuan Indonesia.
- Athallah Muhammad, Abitdavy, Desta Sandya Prasvita, Jl RS Fatmawati Raya, Pd Labu, Kec Cilandak, Kota Depok, and Daerah Khusus Ibukota Jakarta. n.d. 'ANALISIS SENTIMEN PENGGUNA APLIKASI DANA BERDASARKAN ULASAN PADA GOOGLE PLAY MENGGUNAKAN METODE SUPPORT VECTOR MACHINE Prodi S1 Informatika / Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta'.
- Baiti, Faiha Nur, and Irvan Iswandi. 2022. *ANALISIS TRANSAKSI KREDIT SECARA ONLINE PADA APLIKASI AKULAKU DALAM PERSPEKTIF HUKUM POSITIF DAN HUKUM ISLAM (Studi Pada Mahasiswa IAI AL-AZIS)*. Vol. 1.
- Daryfayi, Edyt, Putra Daulay, and Ibnu Asror. n.d. *Sentimen Analisis Pada Ulasan Google Play Store Menggunakan Metode Naïve Bayes*.
- Djamzuri, Muhammad Irfan, and Agung Putra Mulyana. 2022. 'Fenomena Netflix Platform Premium Video Streaming Membangun Kesadaran Cyber Etik Dalam Perspektif Ilmu Komunikasi'. *Jurnal Ilmu Sosial Dan Pendidikan (JISIP)* 6(1):2598–9944. doi: 10.36312/jisip.v6i1.2804/http.
- Fakhri Irfani, Faizal, and Mohamad Triyanto. n.d. *Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine*.



- Fani Al-shufi, Muhamad, and Adhitia Erfina. 2021a. *SENTIMEN ANALISIS MENGENAI APLIKASI STREAMING FILM MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DI PLAYSTORE*. UNIVERSITAS BINA INSAN FAKULTAS KOMPUTER. Vol. 12.
- Fani Al-shufi, Muhamad, and Adhitia Erfina. 2021b. *SENTIMEN ANALISIS MENGENAI APLIKASI STREAMING FILM MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DI PLAYSTORE*. Purba, Mariana, and Inka Rizki Padya. 2023. *ANALISIS SENTIMEN MARKETPLACE DI ERA SOCIETY 5.0 MENGGUNAKAN ALGORITMA NAIVE BAYES*. Vol. 8.
- Gormantara, Alfredo. 2020. *Analisis Sentimen Terhadap New Normal Era Di Indonesia Pada Twitter Menggunakan Metode Support Vector Machine*. Sudarsono, Bernadus Gunawan, Marcell Ignatius Leo, Ali Santoso, and Felix Hendrawan. 2021. 'ANALISIS DATA MINING DATA NETFLIX MENGGUNAKAN APLIKASI RAPID MINER'. *JBASE - Journal of Business and Audit Information Systems* 4(1). doi: 10.30813/jbase.v4i1.2729.
- Indriani, Agnest, Cecep Hermana, Fakultas Ekonomi, and Singaperbangsa Karawang. 2023. 'Analisis Harga Pada Minat Konsumen Dalam Berlangganan Netflix Pasca Pandemi'. *JAMBURA* 6(1). Tanggraeni, Artanti Inez, and Melkior N. N. Sitokdana. 2022. 'Analisis Sentimen Aplikasi E-Government Pada Google Play Menggunakan Algoritma Naïve Bayes'. 9(2):785–95.
- Iqbal Ahmadi, Muhammad, Fuji Apriani, Mia Kurniasari, Siti Handayani, Dudih Gustian, Jl Raya Cibatua Cisaat No, Cibolang Kaler, and Jawa Barat. n.d. 'SENTIMENT ANALYSIS ONLINE SHOP ON THE PLAY STORE USING METHOD SUPPORT VECTOR MACHINE (SVM)'. *Seminar Nasional Informatika* 2020. Wahyudi, Rizki, Gilang Kusumawardhana, Amikom Purwokerto, Jl Letjend, Pol Soemarto, Kec Purwanegara, Tim Purwokerto, and Kabupaten Banyumas. 2021. 'Analisis Sentimen Pada Review Aplikasi Grab Di Google Play Store Menggunakan Support Vector Machine'. *JURNAL INFORMATIKA* 8(2).
- Laia, Erlina, and M. Yamin. 2023. 'Penerapan Algoritma Naïve Bayes Dalam Menganalisis Sentimen Pada Review Pengguna E-Commerce'. *Media Online* 4(1):305–16. doi: 10.30865/klik.v4i1.1186.
- Nur Alamsyah, M. 2020. *PENERAPAN ALGORITMA NAIVE BAYES UNTUK PREDIKSI PENERIMAAN MAHASISWA BARU STUDI KASUS*