

Analisis Sentimen Terhadap Ulasan Aplikasi Disney+ Hotstar Pada Google Playstore Menggunakan Metode Naïve Bayes

Reza Al Arsad, Erizal*

Fakultas Teknologi industri dan Informatika, Teknik Informatika, Universitas Muhammadiyah Prof. Dr. Hamka, Jakarta, Indonesia

Email: ¹alarsad_reza@email.com, ²*erizal@uhamka.ac.id

Email Penulis Korespondensi: erizal@uhamka.ac.id

Submitted: 07/01/2025; Accepted: 26/02/2025; Published: 01/03/2025

Abstrak—Teknologi di Indonesia sudah melesat pesat membuat banyak perubahan seluruh aspek kehidupan, salah satunya adalah aspek streaming online yaitu aplikasi Disney+ Hotstar. Kini Disney+ Hotstar tersedia di tablet, TV cerdas, komputer, maupun ponsel cerdas diakses berbagai tempat dan waktu. Disney+ Hotstar memiliki ribuan jam berbagai film Pixar, Marvel, serta film eksklusif serial Indonesia maupun berbagai negara. Meskipun Disney+ Hotstar memiliki beragam film dan fitur yang menarik, tidak menjamin pengguna merasa puas menggunakan aplikasinya. Dikarenakan pengguna memiliki beragam pendapat dan penilaian, poin ini terlihat dari ulasan pengguna yang tersedia di Google Playstore. Tujuan utama dari studi ini guna mengetahui penilaian atau sentimen dari tinjauan pengguna terhadap aplikasi Disney+ Hotstar dengan menganalisisnya. Teknik yang dipakai menggunakan algoritma Naïve Bayes. Sebanyak 1000 data ulasan didapatkan pada 28 Desember 2024 dari Google Playstore melalui Google Colab, kemudian diproses menggunakan RapidMiner. Dataset melalui tahap cleaning dan preprocessing menjadi 873 data ulasan. Terdapat ulasan baik sebanyak 128 data dan ulasan buruk sebanyak 745 data. Pembobotan TF-IDF dilakukan sebelum melakukan klasifikasi menggunakan 873 dataset. Tahap klasifikasi menggunakan sistem cross-validation serta menerapkan pendekatan Naïve Bayes. Pengujian dari studi ini mengungkapkan hasil akurasi dari algoritma Naïve Bayes sebesar 76,06%, presisi sebesar 34,12%, dan recall sebesar 67,97%.

Kata Kunci: Analisis Sentimen; Naïve Bayes; Disney+ Hotstar; Google Playstore; Ulasan

Abstract—Technology in Indonesia has advanced rapidly, making many changes in all aspects of life, one of which is the online streaming aspect, namely the Disney+ Hotstar application. Now Disney+ Hotstar is available on tablets, smart TVs, computers, and smartphones accessed from various places and times. Disney+ Hotstar has thousands of hours of various Pixar, Marvel films, as well as exclusive Indonesian and various countries' series. Although Disney+ Hotstar has a variety of interesting films and features, it does not guarantee that users are satisfied using the application. Because users have different opinions and assessments, this point can be seen from user reviews available on the Google Playstore. The main purpose of this study was to determine the assessment or sentiment of user reviews of the Disney+ Hotstar application by analyzing it. The technique used uses the Naïve Bayes algorithm. A total of 1000 review data were obtained on December 28, 2024 from the Google Playstore via Google Colab, then processed using RapidMiner. The dataset went through the cleaning and preprocessing stages to become 873 review data. There were 128 good reviews and 745 bad reviews. TF-IDF weighting was performed before classification using 873 datasets. The classification stage used a cross-validation system and applied the Naïve Bayes approach. Testing from this study revealed the accuracy results of the Naïve Bayes algorithm of 76.06%, precision of 34.12%, and recall of 67.97%.

Keywords: Sentiment Analysis; Naïve Bayes; Disney+ Hotstar; Google Playstore; Review

1. PENDAHULUAN

Teknologi di Indonesia meluas dengan pesat yang membuat banyaknya perubahan di seluruh aspek kehidupan. Dengan meluasnya perkembangan teknologi, terdapat berbagai cara dalam menyenangkan diri sendiri dengan melakukan langganan dalam aplikasi streaming online [1]. Layanan streaming online menjadi kegiatan hiburan bagi masyarakat untuk ditonton yang berguna menghilangkan rasa jenuh pada diri sendiri [2]. Banyak dari masyarakat di Indonesia terutama pengguna internet menggunakan aplikasi layanan internetnya untuk meningkatkan kegembiraan dalam menonton hiburan film maupun tv show [3]. Tersedia di beragam perangkat dalam menontonnya, mencakup tablet, TV cerdas, komputer, maupun ponsel cerdas yang dapat diakses diberbagai tempat dan waktu [4], serta layanan hiburan yang memiliki sistem berbayar mempromosikan berbagai bahasa, genre, dan pembuatan film yang dikembangkan melalui berbagai negara diseluruh penjuru dunia [2]. Pemberian rekomendasi film maupun tv show memberikan dampak sangat menarik serta seru yang memberikan pengguna dalam rasa ketagihan serta ingin menontonnya lagi [2]. Adapun dengan keberadaan dari aplikasi streaming online seperti Disney+ Hotstar yang terwujud berkat dukungan dari layanan teknologi digital yang telah memperlancar pengguna dalam mengaksesnya [5]. Aplikasi Disney+ Hotstar sendiri memiliki ribuan jam yang dapat menghibur dengan berbagai film Pixar, Marvel, serta tersedia film eksklusif yang meyiarkan film dari serial Indonesia maupun serial dari beberapa negara [1]. Tidak hanya menyajikan film dan serial yang menarik, tetapi berbagai fitur dan inovatifnya membuat antarmuka dari pengguna menjadi ramah dan sistem fungsi dari pencarian yang futuristik, serta Disney+ Hotstar juga tidak lupa meninggalkan terkait peningkatan pengalaman menonton kepada pengguna [5]. Tetapi tidak menutup kemungkinan bahwa semua pengguna akan merasa puas dengan pengalaman menggunakan Disney+ Hotstar. Namun dalam penerapannya, efek negatif yang diberikan pengguna pun tetap tidak memungkinkan untuk dapat di elakkan [6].

Layanan streaming online Disney+ Hotstar awalnya merupakan hanya sebuah layanan video yang berasal dari India, PT The Walt Disney Company yang berada dibawah kepemilikan Star India [7]. Pada bulan April 2020

diakuisisi Hotstar oleh layanan Disney+ [8]. Aplikasi ini tersedia di Google Playstore, Disney+ Hotstar menjadi salah satu platform streaming online yang paling sering diunduh mencapai 100 juta+ pengguna yang mengunduh pada Google Play Global dan 10 juta+ pengguna yang mengunduh pada Google Play Indonesia. Google Playstore menjadi layanan penyediaan digital yang dikembangkan oleh Google, ulasan menjadi salah satu fitur dari Google Playstore yang memungkinkan pengguna untuk melakukan penilaian terkait dari tingkat kepuasan pengguna itu sendiri yang menggunakan aplikasinya [9]. Wadah untuk memperoleh beragam aplikasi maupun konten digital ada didalam Google Playstore yang mana pengguna dapat menggunakannya melalui ponsel pintar Android, terdiri dari elektronik-book, permainan, maupun film dan juga lainnya; dengan mengakses secara berbayar maupun yang secara gratis juga tersedia [10]. Disney+ Hotstar memiliki beragam macam film serta kemudahan fitur bagi penggunanya, akan tetapi beberapa pengguna masyarakat tidak sepenuhnya menyukai dari keseluruhan film tersebut, namun terdapat juga pengguna yang tidak begitu nyaman dengan fitur terkait dari tampilan penggunaan aplikasi Disney+ Hotstar itu sendiri. Fitur penilaian serta ulasan menjadi fitur yang sangat penting pada Google Playstore. Kemampuan dari fitur ini dapat memberikan dampak yang sangat berharga dalam membuat pengguna melakukan penilaian terkait aplikasi yang digunakan dalam seberapa baiknya aplikasi tersebut [11].

Analisis sentimen, yang diketahui secara lain dengan sebutan penelusuran pendapat yakni kajian yang mengevaluasi pandangan, pikiran, maupun impresi yang dimiliki seorang individu dalam layanan teks data, serta mengklasifikasikan polaritas sentimen baik dan buruknya [12]. Menentukan ulasan maupun pendapat yang memiliki kelas positif atau negatif menjadi peran krusial dalam analisis sentimen dengan penggunaan metode klasifikasi Naïve Bayes dalam analisis sentimen memiliki peran krusial. Dengan adanya informasi berikut dapat menjadikan acuan dalam meningkatkan kualitas dari layanan atau aplikasi [13]. Pendekatan yang biasanya digunakan dalam analisis sentimen dengan beberapa penggunaan seperti machine learning salah satunya [14]. Dengan diterapkannya metode algoritma Naïve Bayes dalam studi ini, disebabkan keahliannya saat menjalankan analisa untuk menentukan kelas positif maupun negatif menjadi kecepatannya serta kemudahannya [15]. yang mana dapat menghemat waktu serta mengurangi upaya yang dibutuhkan untuk mengklasifikasi ulasan [14].

Pada penelitian sebelumnya yang dilakukan oleh Gilbert Darmawan, dan timnya pada tahun 2023 meneliti mengenai analisis sentimen berdasarkan Ulasan Pengguna aplikasi MyPertamina pada Google Playstore menggunakan metode Naïve Bayes, menunjukkan tingkat dari akurasi sebanyak 91%, presisi sebesar 92%, dan 100% untuk tingkat recallnya yang tergolong negatif dalam 3948 data yang digunakan [16]. Menurut studi peneliti yang dilakukan oleh Hilmy Zhafran Muflih, Allif Rizki Abdillah, dan Firman Noor Hasan pada tahun 2023 menganalisis penelitian terkait analisis sentimen terhadap Ulasan Pengguna Aplikasi Ajaib dengan menggunakan Metode Naïve Bayes, dari analisis tersebut mendapatkan 76,67% tingkat dari akurasi dengan menggunakan data sebanyak 500 data [17]. Berdasarkan penelitian tahun 2024 yang dilakukan oleh Kevin.K, Margaret Enjeli, dan Andri Wijaya dalam penelitian “Analisis Sentimen Penggunaan Aplikasi Kinemaster Menggunakan Metode *Naïve Bayes*”, perbandingan antara data *training* dan data *testing* memperoleh tingkat akurasi sebesar 85%, presisi sebanyak 82%, dan recall dengan 74% [18]. Menggunakan 500 data ulasan, Rizqi Rizaldi dan Riska Aryanti pada tahun 2024 melakukan penelitian terhadap aplikasi Indodana yang memberikan hasil dengan menggunakan *Naïve Bayes* mendapatkan nilai akurasi sebesar 86% [19]. Di sisi lain, penelitian yang dikembangkan oleh Muhammad Irfan dan Erizal melakukan perbandingan antara algoritma *Naïve Bayes* dengan K-Nearest Neighbor pada aplikasi Indrive berhasil mendapatkan nilai untuk algoritma *Naïve Bayes* dari tingkat akurasi 97,50%, presisi sebanyak 92,71%, dan recall sebanyak 100%. Sementara itu untuk algoritma K-NN mendapatkan nilai dengan tingkat akurasi sebesar 83,21%, presisi 85%, dan recal sebesar 57,30% [20].

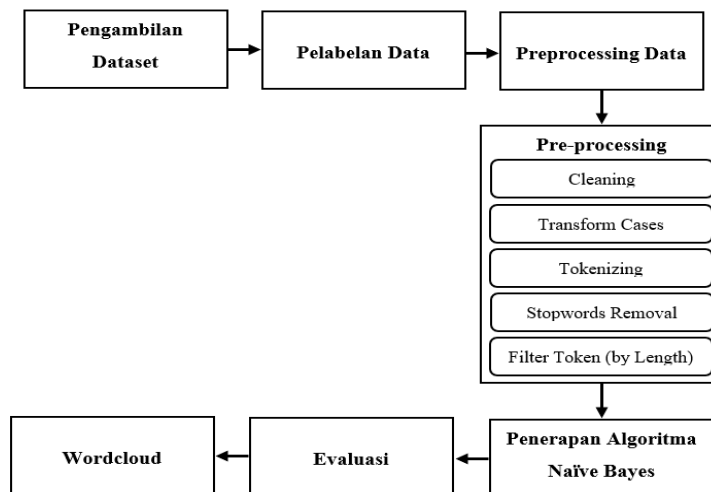
Dengan maksud untuk menegaskan bahwa layanan streaming dari aplikasi Disney+ Hotstar guna bisa menghasilkan efek bahagia dari tingkat kelas positif dan mendapatkan manfaat yang ideal, maka dibutuhkan penelitian lebih lanjut dengan contoh penelitian yang dilakukan oleh Niken Febrinikmah Siharta pada tahun 2024 melakukan penelitian terhadap aplikasi Disney+ Hotstar dengan menggunakan algoritma *Naïve Bayes* dalam penggunaannya teknik scraping data hingga evaluasi tetap dalam penggunaan python dan tools melalui Google Colab, dari eksperimen tersebut memberikan hasil dari tingkat akurasi sebesar 81,6%, presisi 58%, dan keberhasilan mengenai kembali sebesar 83%, serta mendapatkan hasil dari f1-score 68% [7].

Bedasarkan pemaparan dari masalah sebelumnya, Penelitian ini akan memfokuskan pada pengelompokan respon baik maupun buruk dengan menggunakan algoritma *Naïve Bayes*. Pemilihan pendekatan metode ini dilakukan sebab keahliannya dalam menjalankan analisis secara cekat dan dibuktikan oleh Muhammad Irfan dalam penelitian dengan membandingkan metode *Naïve Bayes* dengan metode K-Nearest Neighbor (K-NN), yang mampu menghemat waktu dan tenaga yang diperlukan dalam mengelompokan ulasan. Tujuan utama penelitian ini ialah untuk memahami penilaian atau sentimen dalam ulasan yang baik maupun buruk, merupakan poin utama bagi pencipta aplikasi Disney+ Hotstar dalam meningkatkan performa serta keuntungan untuk pengguna dari aplikasi.

2. METODOLOGI PENELITIAN

Metode yang diuji oleh peneliti dalam tahap penelitian ini dengan menerapkan *Naïve Bayes* yang akan mengidentifikasi sentimen dari ulasan aplikasi melalui hasil review dari segmen ulasan pada *google playstore* dan mengimplementasikan pada sistem perangkat *RapidMiner* untuk mendapatkan hasil prediksi sentimen, yang mana

dapat memudahkan dalam mengenali kelas positif atau negatif dari hasil sentimen. Pada Gambar 1 menampilkan proses dari alur perancangan yang peneliti lakukan.



Gambar 1. Tahap Perancangan

2.1 Pengambilan Dataset

Pengambilan data ulasan yang didapatkan melalui tahapan web scraping yang diimplementasikan menggunakan Google Colab. Web *Scraping* adalah metode pengambilan atau pengumpulan data dan digunakan untuk mengolah informasi dari jumlah besar data yang diambil dari web. Data yang telah dipilih akan diubah menjadi tabel yang serupa dengan *Spreadsheet*, mirip dengan pengumpulan data di Microsoft Excel dalam format CSV [21]. Seluruh hasil pengambilan data dari proses web scraping yang akan dikumpulkan dengan format dokumen CSV untuk digunakan pada tahapan berikutnya.

2.2 Pelabelan Data

Tahap pelabelan data merupakan bagian dari mengelompokkan dua kelas data: negatif dan positif, pengelompokan kelas baik diterapkan untuk memvisualisasikan ekspresi baik seperti ceria, bahagia, rileks, dan ekspresi yang berkaitan dengan rasa bahagia. Namun sebaliknya untuk kelas negatif diimplementasikan untuk emosi marah, takut, sedih, tertekan, dan emosi yang dapat menyebabkan merasakan penderitaan. Proses pembagian ini bersifat subjektif dan dilakukan secara manual [17].

2.3 Preprocessing Data

Membersihkan dataset adalah langkah awal dalam proses ini [22]. Untuk mengoptimalkan proses klasifikasi, setiap teks ulasan yang mengandung angka, emoji, tanda baca, atau spasi yang tidak perlu dibersihkan. Selain itu, bahasa ulasan yang berulang-ulang juga dibersihkan [23]. Pada tahapan preprocessing dilakukan 5 langkah, yaitu *cleaning*, *transform cases*, *tokenize*, *stopwords removal*, dan *filter token (by length)*.

2.4 Term Frequency – Inverse Document Frequency (TF-IDF)

Memanfaatkan metode tahapan bahasa alami yang akan dikenal pada tahap ini dengan nama Frekuensi dokumen terbalik atau TF-IDF, kopos menjadi nama lain dari pengukuran tingkat pentingnya suatu kata yang berada dalam dokumen atau kumpulan dari dokumen yang merupakan tujuan dari TF-IDF. Teknik ini menghubungkan antara dua komponen *primer*, yang mana berupa Term Frequency (TF) serta Inverse Document Frequency (IDF). Membandingkan frekuensi kemunculan suatu bahasa dengan total frekuensi kata yang ada didalam berkas merupakan teknik dari TF (Term Frequency) dalam menghitung frekuensi kemunculan suatu kata dalam dokumen[24]. Sebaliknya, IDF mengukur tingkat dari kelangkaan atau keunikan suatu istilah dalam seluruh korpus dengan membandingkan antara logaritma dari rasio jumlah keseluruhan dokumen terhadap jumlah dokumen yang memuat kata tersebut [25].

2.5 Penerapan Algoritma Naïve Bayes Classifier

Pada Penelitian ini *Naïve Bayes Classifier* merujuk pada sebuah pendekatan dengan teknik klasifikasi didasarkan pada prinsip teorema Bayes. Pendekatan ini memanfaatkan probabilitas masa depan yang didasarkan pada data dari pengalaman sebelumnya, sehingga sering disebut sebagai Teorema Bayes [26]. *Naïve Bayes Classifier* juga dikenal sebagai salah satu metode atau alat klasifikasi yang sangat efisien [27]. Setiap ulasan yang melalui tahapan preprocessing akan diberikan nilai bobot dari pengimplementasian algoritma *Naïve Bayes Classifier*. Seiring dengan meningkatnya dari frekuensi kemunculan kata-kata terkait yang ditunjuk untuk pembobotan, maka angka pembobotannya akan mengalami peningkatan [11].

2.6 Evaluasi

Evaluasi merupakan bagian dari proses peninjauan ulang untuk memverifikasi keakuratan dan ketepatan hasil uji. Verifikasi ketepatan hasil uji bertujuan untuk mengidentifikasi hasil uji yang optimal. Evaluasi ketepatan model dilakukan melalui pemanfaatan matriks kebingungan (*confusion matrix*). Penilaian akurasi ini mencakup evaluasi melalui akurasi, presisi, dan recall [28]. Perumusan dalam menghitung *confusion matrix* untuk menguji dari tingkat akurasinya berikut ini:

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{1}$$

$$Precision = \frac{TP}{(TP+FP)} \tag{2}$$

$$Recall = \frac{TP}{(TP+FN)} \tag{3}$$

Adapun berikut ini terdapat 4 sebutan dari representasi hasil pada tahapan klasifikasi dengan perumusan yang mencakup *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN). [29]. Berikut Gambar 2 bentuk tampilan dari Akurasi, Presisi, dan Recall.

	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

Gambar 2. Tampilan representasi *confusion matrix*

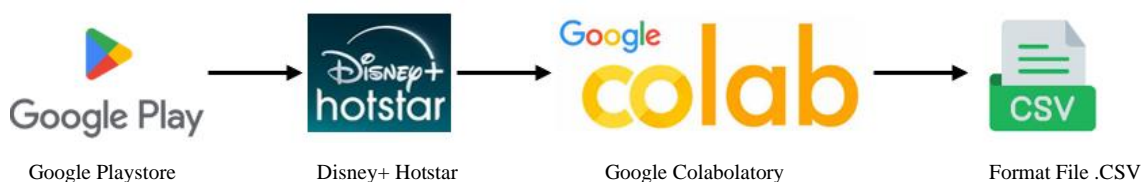
2.7 Wordcloud

Peneliti menggunakan visualisasi data dalam bentuk wordcloud dalam penelitian ini. Wordcloud merupakan teknik visualisasi yang menampilkan kata-kata yang paling menonjol berdasarkan keseringan munculnya dari suatu kata dalam teks. Wordcloud digunakan dalam analisis teks memberikan peneliti sebuah representasi cepat mengenai pokok pembahasan dari teks yang dianalisis [30].

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Peneliti memperoleh data ulasan memakai proses web scraping yang dilakukan menggunakan *Google Colab*. Tahap web scraping aplikasi streaming Google Playstore Disney+ Hotstar ini menentukan jumlah data yang dikumpulkan sebanyak 1000 komentar/ulasan yang didapatkan pada tanggal 28 Desember 2024. Peneliti harus menginput kode pemrograman python menggunakan *Google Colab* untuk mendapatkan ulasan pada Google Playstore.



Gambar 3. Proses perolehan ulasan data

Pada Gambar 3 peneliti melakukan tahapan perolehan ulasan data yang dilakukan teknik web scraping dengan dimulai dari mengakses Google Playstore, dilanjutkan mendapatkan link pada aplikasi Disney+ Hotstar, kemudian menjalankan Google Colab serta memasukan link aplikasi Disney+ Hotstar pada coding pemrograman untuk memperoleh data melalui scraping yang mana akan disimpan dalam format file CSV. Maka pada gambar 4 bisa dilihat hasil dari pengumpulan dengan menggunakan teknik scrapping menggunakan Google Colab.

```

from google_play_scraper import Sort, reviews

result, continuation_token = reviews(
    'in.startv.hotstar.dplus',
    lang='id', # defaults to 'en'
    country='id', # defaults to 'us'
    sort=Sort.NEWEST, # defaults to Sort.MOST_
    count=1000, # defaults to 100
    filter_score_with=None # defaults to None
)
  
```

Gambar 4. Crawling data Google Colab

Pada Gambar 4 peneliti memasukan coding “*in.startv.hotstar.dplus*” untuk mendapatkan data ulasan dari aplikasi Disney+ Hotstar pada aplikasi versi Indonesia. Peneliti memasukan jumlah data yang akan didapatkan dengan memasukan kode “*count=1000*”. Seluruh hasil pengumpulan data dari proses web scraping disimpan dalam format file .CSV untuk digunakan pada tahapan berikutnya.

3.2 Pelabelan Data

Proses Dalam tahap ini, proses pelabelan merupakan langkah di mana dataset diberi label untuk kelas positif atau negatif. Data yang telah dihasilkan melalui proses *scraping* dan berbentuk csv yang akan diberi label secara manual. Pendekatan ini memungkinkan untuk menandai objek data terdapat dalam ulasan, memudahkan proses identifikasi sentimen dalam dataset.

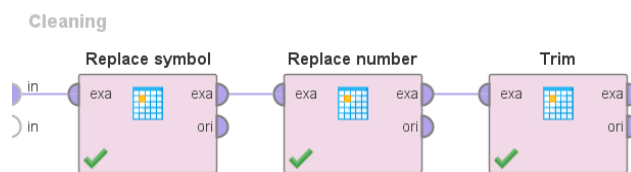
Tabel 1. Tahapan proses *labeling*

Label	Ulasan
Positif	Bagus banget bisa nonton marvel sepuasnya
Negatif	Aplikasi apa ini mau login aja gak bisa
Positif	Semoga aplikasinya bagus dan bisa digunakan saat dimanapun.
Negatif	Aplikasi gak jelas, udah langganan aturan 1 bulan full, tpi ini belum 25 hari aja masa udah abis masa berlaku nya kocak!!

Terdapat 4 contoh hasil pelabelan yang berupa 2 positif dan 2 negatif dalam Tabel 1.

3.3 Preprocessing Data

Proses selanjutnya melakukan preprocessing dataset pada ulasan aplikasi *Disney+ Hotstar* yang telah diberikan label dan masih tidak teratur dengan menggunakan *RapidMiner*. Pada Gambar 5 tahap ini dimulai dengan proses *Cleaning*.



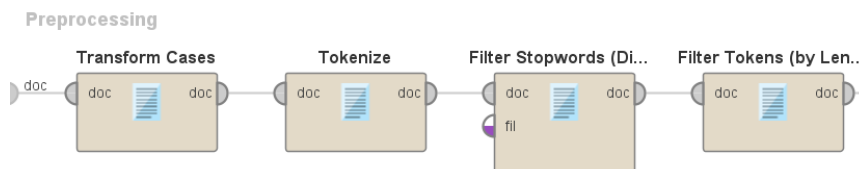
Gambar 5. Penggunaan operator dalam *cleaning* data

Pada tahap *Cleaning* merupakan proses yang dilakukan untuk menghapus data duplikat pada ulasan aplikasi *Disney+ Hotstar* di Google Play Store yang sama atau identik, sekaligus membersihkan kalimat-kalimat yang tidak relevan dan tidak memiliki nilai. Pada proses ini dilakukan pembersihan dataset seperti tanda baca, simbol-simbol, dan emoticon yang terletak pada dataset. Adapun Tabel 2 contoh dari tahap ulasan sebelum dan sesudah dilakukan tahapan *cleaning*.

Tabel 2. Tahapan *Cleaning*

Sebelum <i>Cleaning</i>	Sesudah <i>Cleaning</i>
Paket tahunan nya, melejit ga pake rem harga naik nya. Padahal sudah berlangganan dari tahun sebelumnya. Harga nya dibanding dengan film2 nya kurang berimbang.	Paket tahunan nya melejit ga pake rem harga naik nya Padahal sudah berlangganan dari tahun sebelum nya Harga nya dibanding dengan film nya kurang berimbang
Aplikasi nya bagus tapi knp layar nya hitam saat mau nonton diperlihatkan hanya Subtitle nya aku Bingung!!	Aplikasi nya bagus tapi knp layarnya hitam saat mau nonton diperlihatkan hanya Subtitle nya aku Bingung

Pada Tabel 2 proses ini membersihkan tanda, angka, dan simbol seperti: (, (.), (2), dan (!). Pada gambar 6 menunjukkan tahapan proses *preprocessing* data pada penelitian ini yang menggunakan beberapa operator dari *RapidMiner*. Berikut dari operator yang digunakan yaitu: *Transform Case*, *Tokenize*, *Stopword*, dan *Filter Token(by Length)*.



Gambar 6. Penggunaan operator dalam preprocessing data

Selanjutnya tahap *Transform Cases/Case Folding* merupakan proses dimana dataset yang sudah dibersihkan kemudian diubah menjadi font kecil (lowercase), sehingga memudahkan sistem untuk mengenali huruf yang serupa

dengan makna yang serupa. Berikut tahapan ulasan yang sebelum dilakukan dan sesudah dilakukan tahap *cleaning* dalam Tabel 3 contohnya.

Tabel 3. Tahapan *Transform Cases*

Sebelum <i>Transform Cases</i>	Sesudah <i>Transform Cases</i>
Paket tahunan nya melejit ga pake rem harga nya naik Padahal sudah berlangganan dari tahun sebelum nya Harga nya dibanding dengan film nya kurang berimbang	paket tahunan nya melejit ga pake rem harga naik nya padahal sudah berlangganan dari tahun sebelum nya harga nya dibanding dengan film nya kurang berimbang
Aplikasi nya bagus tapi knp layar nya hitam saat mau nonton diperlihatkan hanya Subtitle nya aku Bingung	aplikasi nya bagus tapi knp layarnya hitam saat mau nonton diperlihatkan hanya subtitle nya aku bingung

Pada proses Tabel 3 berhasil membuat semua font menjadi kecil (lowercase). Berikutnya tahap *Tokenize* merupakan tahapan membagi teks atau menguraikan setiap kata, yang merupakan proses pemisahan kalimat menjadi unit-unit kata yang terpisah serta memudahkan dalam proses analisis yang merupakan proses dari tokenizing. Prosesnya memberikan token dengan cara membagi kalimat data menjadi beberapa potongan kecil kata yang disebut dengan nama token. Berikut tabel 4 contoh dari tahapan ulasan yang sebelum dan sesudah dilakukan tahap *Tokenize*.

Tabel 4. Tahapan *Tokenize*

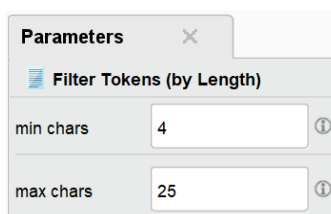
Sebelum <i>Tokenize</i>	Sesudah <i>Tokenize</i>
paket tahunan nya melejit ga pake rem harga naiknya padahal sudah berlangganan dari tahun sebelum nya harga nya dibanding dengan film nya kurang berimbang	“paket” “tahunan” “nya” “melejit” “ga” “pake” “rem” “harga” “naik” “nya” “padahal” “sudah” “berlangganan” “dari” “tahun” “sebelum” “nya” “harga” “nya” “dibanding” “dengan” “film” “nya” “kurang” “berimbang”
aplikasi nya bagus tapi knp layar nya hitam saat mau nonton diperlihatkan hanya subtitle nya aku bingung	“aplikasi” “nya” “bagus” “tapi” “knp” “layar” “nya” “hitam” “saat” “mau” “nonton” “diperlihatkan” “hanya” “subtitle” “nya” “indonesia” “juga” “banyak” “film” “marvel” “lengkap” “aku” “bingung”

Pada tabel 4 berhasil memisahkan keseluruhan kalimat menjadi satu kata untuk memudahkan dalam melakukan proses selanjutnya. Tahap *Stopwords Removal* merupakan proses pembuangan *stopwords* atau kata-kata tidak penting/dianggap kurang bermakna dari hasil sentimen token, seperti “di”, “kok”, “dan”, “yg”, “pas”, “pada”, “kek”, dan hal-hal yang tidak terkait dengan penelitian ini. Berikut Tabel 5 contoh dari proses ulasan sebelum dan sesudah dilakukan tahap *Stopwords Removal*.

Tabel 5. Tahapan *Stopwords Removal*

Sebelum <i>Stopwords Removal</i>	Sesudah <i>Stopwords Removal</i>
“paket” “tahunan” “nya” “melejit” “ga” “pake” “rem” “harga” “naik” “nya” “padahal” “sudah” “berlangganan” “dari” “tahun” “sebelum” “nya” “harga” “nya” “dibanding” “dengan” “film” “nya” “kurang” “berimbang”	“paket” “tahunan” “melejit” “nya” “ga” “pake” “rem” “harga” “naik” “nya” “berlangganan” “tahun” “sebelum” “nya” “harga” “nya” “dibanding” “dengan” “film” “kurang” “berimbang”
“aplikasi” “nya” “bagus” “tapi” “knp” “layar” “nya” “hitam” “saat” “mau” “nonton” “diperlihatkan” “hanya” “subtitle” “nya” “indonesia” “juga” “banyak” “film” “marvel” “lengkap” “aku” “bingung”	“aplikasi” “nya” “bagus” “layar” “hitam” “mau” “nonton” “diperlihatkan” “subtitle” “nya” “indonesia” “banyak” “film” “marvel” “lengkap” “aku” “bingung”

Pada Tabel 5, sudah membersihkan beberapa kata yang tidak penting/signifikan dalam penelitian ini dengan menggunakan *stopword removal*. Pada Gambar 7 memperlihatkan operator parameters dengan minimum 4 dan maksimum 25 karakter dalam membatasi limit huruf dari *filter token (by Length)* yang akan dilakukan untuk proses selanjutnya dalam Tabel 6.



Gambar 7. Tampilan operator parameters limitasi karakter

Pada Tabel 6 melakukan, tahap *Filter Token (by Length)* bertujuan untuk menentukan jumlah karakter saat proses penyeleksian token yang sedang dilakukan berdasarkan panjangnya kalimat data. Berikut ini contoh dari proses ulasan sebelum dan sesudah dilakukan tahap *Filter Token (by Length)* pada Tabel 6.

Tabel 6. Tahapan *Filter Token (by Length)*

Sebelum <i>Filter Token (by Length)</i>	Sesudah <i>Filter Token (by Length)</i>
“paket” “tahunan” “melejit” “nya” “ga” “pake”	“paket” “tahunan” “melejit” “pake” “harga” “naik”
“rem” “harga” “naik” “nya” “berlangganan” “dari”	“berlangganan” “tahun” “sebelum” “harga” “dibanding”
“tahun” “sebelum” “nya” “harga” “nya”	“dengan” “film” “kurang” “berimbang”
“dibanding” “dengan” “film” “kurang”	“aplikasi” “bagus” “layar” “hitam” “nonton” “diperlihatkan”
“berimbang”	“subtitle” “indonesia” “banyak” “film” “marvel” “lengkap”
“aplikasi” “nya” “bagus” “layar” “hitam” “mau”	“bingung”
“nonton” “diperlihatkan” “subtitle” “nya”	
“indonesia” “banyak” “film” “marvel” “lengkap”	
“aku” “bingung”	

Memperlihatkan kesimpulan dari *filter token (by Length)* dalam Tabel 6, oleh karena itu kata-kata yang kurang dari 4 karakter maka akan terhapus seperti: “nya”, “ga”, “rem”, dan “aku”.

Setelah melalui tahap preprocessing, data yang awalnya pada penelitian ini sebanyak 1000 dataset dan berhasil dibersihkan menggunakan preprocessing, maka dataset menjadi sebanyak 873 dataset.

3.4 Pembobotan TF-IDF

Fase berikut ini dilakukannya tahap pembobotan TF-IDF memakai data yang telah melalui tahapan preprocessing. Menghitung Term Frequency (TF) merupakan langkah pertama yang dilakukan dari metode ini untuk menilai seberapa sering kemunculan kata yang ada didalam dokumen. Dalam kaidah tertentu, TF menunjukkan frekuensi kemunculan dari sebuah kata pada dokumen tertentu. Kemudian dilakukannya perhitungan pada Inverse Document Frequency (IDF).

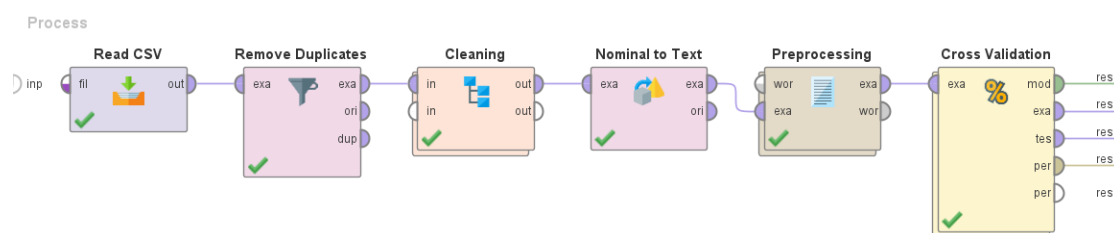
Tabel 7. Tampilan proses pembobotan dari TF-IDF

Teks	aplikasi	enak	film	nonton	tolol
bagus enak filem	0	0.597	0	0	0
aplikasi tolol	0.543	0	0	0	0.840
pilihan film ditambahin	0	0	0.423	0	0
payah aplikasi masuk kode pakai wifi	0.304	0	0	0	0
lawak gabisa login tolol	0	0	0	0	0.518
nonton asyik	0	0	0	0.546	0
aplikasi sampah perpanjang otomatis motong dana	0.256	0	0	0	0

Tabel 7 mempresentasikan hasil penilaian tahap pembobotan TF-IDF pada RapidMiner, kemunculan dalam dokumen dari setiap kata-kata diberi nilai bobot yang berdasarkan frekuensinya.

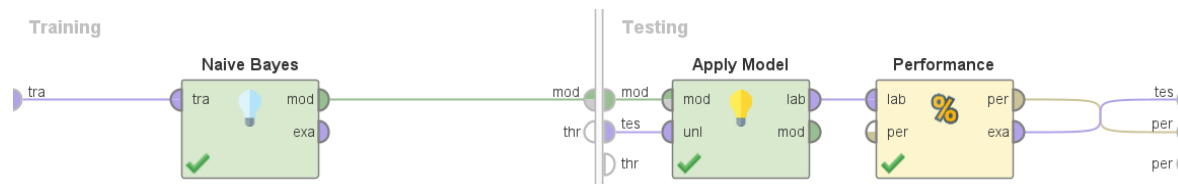
3.5 Algoritma Naïve Bayes Classifier

Setelah sebelumnya melakukan tahapan *preprocessing*, maka dilanjutkan ke tahapan berikutnya pelaksanaan dari algoritma Naïve Bayes, yang mana menggunakan metode Naïve Bayes data tersebut akan diklasifikasikan. Pada tahap uji ini peneliti melakukan *cross-validation* bertujuan untuk mengevaluasi dari tingkat keakurasian pada dataset dan algoritma *Naïve Bayes* menerapkan hasil klasifikasi. Operator *cross-validation* sangat dibutuhkan dalam melakukan tahap penerapan dengan pendekatan *k-fold cross-validation* dengan k = 10, memecah data menjadi 10 subset dengan sama rata merupakan tujuan dari teknik *k-fold cross-validation*.



Gambar 8. Tahapan klasifikasi pada *Naïve Bayes*

Menyajikan Gambar 8 dari keseluruhan tahap pembagian data dengan menggunakan algoritma *Naïve Bayes* ini. Tujuan dari mengukur kinerja algoritma dengan menghubungkan semua operator dimulai dari dataset yang sudah dilabeli hingga operator *cross-validation*.



Gambar 9. Tahap penggunaan operator *cross-validation*

Pada Gambar 9 memperlihatkan pengoperasian yang dimanfaatkan dalam *cross-validation* terdapat operator *Naive Bayes* bertujuan dalam mengklasifikasi data, operator *Apply Model* berfungsi dalam mengolah data masukan serta memberikan prediksi output menggunakan teknik klasifikasi, operator terakhir *Performance* berperan dalam menguji dari kinerja algoritma *Naive Bayes* dalam mengklasifikasi data.

3.6 Evaluasi

Setelah melakukan tahap klasifikasi algoritma *Naive Bayes* menggunakan *Rapidminer* maka tahap selanjutnya untuk mengetahui hasil dari pengujian/evaluasi yang memperoleh tingkat dari *accuracy*, *precision*, dan *recall*. Tingkat dari *accuracy* berhasil diperoleh sebesar 76,06%, sedangkan untuk *precision* diperoleh sebesar 34,12%, dan untuk *recall* sebesar 67,97%. Dalam bentuk *confusion matrix* terdapat 87 *True Positive* (TP), 557 *True Negative* (TN), 168 *False Positive* (FP), dan 41 *False Negative* (FN).

accuracy: 76.06% +/- 3.74% (micro average: 76.06%)

	true Negatif	true Positif	class precision
pred. Negatif	577	41	93.37%
pred. Positif	168	87	34.12%
class recall	77.45%	67.97%	

Gambar 10. Tampilan dari capaian pengujian algoritma *Naive Bayes*

Tampilan visualisasi yang didapatkan pada hasil pengujian algoritma *Naive Bayes* menggunakan perangkat *Rapidminer* seperti Gambar 10. Pemaparan dari visualisasi ini mengenai hasil kerja algoritma dalam melakukan klasifikasi data. Berikut uraian dari hasil perhitungan untuk menilai *accuracy*, *precision*, dan *recall* berdasarkan *confusion matrix* dari data yang didapatkan melalui penerapan algoritma *Naive Bayes*:

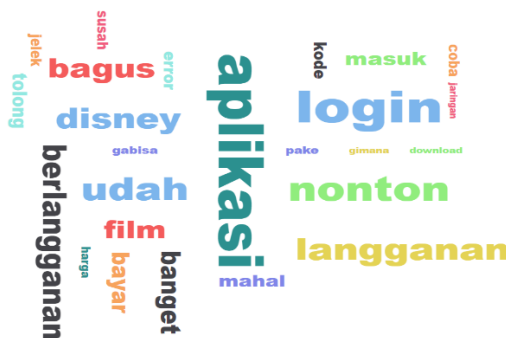
$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} = \frac{(87+577)}{(87+577+168+41)} = \frac{664}{873} = 0,7605 = 76,05\% \text{ atau } 76,06\% \quad (4)$$

$$Precision = \frac{TP}{(TP+FP)} = \frac{87}{(87+168)} = \frac{87}{255} = 0,3412 = 34,12\% \quad (5)$$

$$Recall = \frac{TP}{(TP+FN)} = \frac{87}{(87+41)} = \frac{87}{128} = 0,6797 = 67,97\% \quad (6)$$

3.7 Visualisasi Wordcloud

Wordcloud di tahap ini digunakan untuk memperlihatkan bentuk visualisasi yang mengilustrasikan dari kata-kata yang sering muncul maka mendapatkan bentuk frekuensi yang lebih besar, oleh karena itu sebaliknya kata-kata yang tidak sering muncul mendapatkan bentuk frekuensi lebih kecil.



Gambar 11. Tampilan visualisasi dari hasil *wordcloud*

Representasi dari Gambar 11 merupakan bentuk visual pada aplikasi Disney+ Hotstar, yang menindikasikan bahwa frekuensi kemunculan kata berada di antara 1 hingga 25 yang sering muncul setelah data disusun dari nilai terbesar ke terkecil. Bentuk visualisasi dari *wordcloud* memperlihatkan kata yang frekuensinya paling menonjol yaitu

kata ‘aplikasi’ sebanyak 165 kali, untuk kata kedua yang sering muncul yaitu kata ‘login’ dengan frekuensi sebanyak 146 kali, dan berikutnya frekuensi kata sering muncul terdapat dua kata yaitu ‘nonton’ dengan frekuensi yang muncul sebanyak 116 kali.

4. KESIMPULAN

Dapat diambil kesimpulan dari penelitian diatas memperlihatkan bahwa dataset yang didapatkan pada bulan Desember 2024 sebanyak 1000 dataset ulasan dari aplikasi Disney+ Hotstar pada Google Playstore, adapun teknik dari pengambilan dataset tersebut dengan memanfaatkan teknik web scraping melalui Google Colab. Dataset yang didapatkan dari teknik web scraping yaitu melalui aplikasi Disney+ Hotstar versi Indonesia serta bentuk ulasan yang berbahasa Indonesia. Dataset yang didapatkan memfokuskan untuk mendapatkan data yang terbaru serta dataset tersebut dikonversi menjadi bentuk format file .CSV. Setelah mendapatkan dataset, maka diberikan label dan dilanjutkan ke Rapidminer dengan dimulai dari tahap pembersihan dan *preprocessing* dari yang data awal sebanyak 1000 dan berhasil dibersihkan menjadi 873 data. Tahap pembersihan dan *preprocessing* mencakup beberapa tahapan yang digunakan terdiri dari proses pembersihan/*cleaning* untuk membersihkan data dari berbagai simbol, angka, dan menghapus data duplikat. Setelah itu dilakukan tahap *preprocessing* dengan menggunakan operator *tokenize*, *transform cases*, *stopwords*, *filter token (by Length)* untuk membatasi maksimum dan minimum sebuah karakter. Setelah tahapan *preprocessing* berhasil mendapatkan data sebanyak 873 data. Terdapat ulasan positif dan negatif dari 873 data tersebut, yang mana berisi ulasan positif sebanyak 128 data dan 745 data berisi ulasan negatif. Banyaknya ulasan negatif yang diakibatkan dari keluhan pengguna terhadap sistem login pada aplikasi yang menjadi penyebab pengguna tidak dapat masuk kedalam aplikasinya dikarenakan beberapa faktor salah satunya yaitu tidak mendapatkan kode OTP, disisi lain terdapat pengguna yang merasa senang terhadap aplikasi yang dapat memudahkan pengguna dalam menonton berbagai macam acara yang disukai serta memiliki keunggulan dalam fitur yang baik dalam ulasan positif bagi pengguna. Tahap selanjutnya yaitu klasifikasi, sebelum melakukan tahap klasifikasi dalam pembobotan kata peneliti mengaplikasikan pendekatan TF-IDF yang akan digunakan. Selanjutnya untuk memulai tahap klasifikasi dengan menggunakan algoritma *Naïve Bayes* serta peneliti menerapkan *cross-validation* melalui pengimplementasian teknik *10-fold cross-validation*. Tahap klasifikasi tersebut mendapatkan hasil dengan tingkat akurasi sebesar 76,06%, presisi 34,12%, dan recall sebanyak 67,97%. Adapun saran yang dapat mengembangkan penelitian ini, peneliti menyarankan dengan melakukan algoritma lain, seperti teknik *split ratio* (pembagian data) atau dengan melakukan perbandingan antara *split ratio* dan *cross-validation* yang dapat meningkatkan tingkat akurasi, presisi, maupun recallnya. Untuk penelitian selanjutnya peneliti juga menyarankan untuk melakukan dengan dataset yang lebih banyak agar mendapatkan tingkat akurasi yang lebih.

REFERENCES

- [1] N. Juliandhono and M. P. Berlianto, “Faktor-Faktor Yang Mempengaruhi Perceived Value Dan Implikasinya Kepada Intention To Subscribe Serta Pengaruhnya Terhadap Social Influence Pada Aplikasi Streaming Film Disney Plus Hotstar,” *Jurnal Manajemen Pemasaran*, vol. 16, no. 2, pp. 77–86, 2022, doi: 10.9744/pemasaran.16.2.77–86.
- [2] U. Kulsum, M. Jajuli, and N. Sulistiyowati, “Analisis Sentimen Aplikasi WETV di Google Play Store Menggunakan Algoritma Support Vector Machine,” *Journal of Applied Informatics and Computing*, vol. 6, no. 2, pp. 205–212, Dec. 2022, doi: 10.30871/jaic.v6i2.4802.
- [3] D. Saputra and M. R. Pribadi, “Analisis Sentimen Masyarakat terhadap Layanan Provider Internet di Indonesia menggunakan SVM,” *MDP Student Conference*, vol. 2, no. 1, pp. 32–41, Apr. 2023, doi: 10.35957/mdp-sc.v2i1.4554.
- [4] F. Abrari and I. M. Sukresna, “Pengaruh Self Congruity, Perceived Price, Brand Trust, dan Customer Engagement terhadap Willingness to Continue and Subscribe pada Disney Plus Hotstar,” *Ekonomis: Journal of Economics and Business*, vol. 8, no. 2, p. 1200, Sep. 2024, doi: 10.33087/ekonomis.v8i2.1601.
- [5] S. Jovanka and A. Maulana, “Analisis Kepuasan Pengguna terhadap Layanan Streaming Video dengan Metode E-service quality: Studi Kasus pada Aplikasi Disney Plus,” *SATESI: Jurnal Sains Teknologi dan Sistem Informasi*, vol. 3, no. 1, pp. 45–51, Apr. 2023, doi: 10.54259/satesi.v3i1.2455.
- [6] E. Yuniar, D. S. Utsalinah, and D. Wahyuningsih, “Implementasi Scrapping Data Untuk Sentiment Analysis Pengguna Dompel Digital dengan Menggunakan Algoritma Machine Learning,” *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 2, no. 1, pp. 35–42, Apr. 2022, doi: 10.25008/janitra.v2i1.145.
- [7] N. F. Siharta, M. M. Bhaswara, W. Firmantara, and A. P. Sari, “Analisis Sentimen Ulasan Pengguna Aplikasi Disney+ Hotstar Menggunakan Naive Bayes Classifier,” *Journal of Multidisciplinary Inquiry in Science Technology and Educational Research*, vol. 1, no. 4, pp. 1847–1855, 2024, doi: 10.32672/mister.v1i4.2120.
- [8] A. H. Nugraha, “Analisis Perbandingan User Experience (UX) Pada Aplikasi Netflix Dengan Disney+ Hotstar Menggunakan Metode User Experience Questionnaire (UEQ),” *Jurnal Penelitian Teknologi Informasi dan Sains*, vol. 2, no. 2, pp. 100–114, Jun. 2024, doi: 10.54066/jptis.v2i2.1944.
- [9] S. Fransiska and A. Irham Gufroni, “Sentiment Analysis Provider by.U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method,” *Scientific Journal of Informatics*, vol. 7, no. 2, pp. 2407–2658, 2020, doi: 10.15294/sji.v7i2.25596.
- [10] P. Aditiya, U. Enri, and I. Maulana, “Analisis Sentimen Ulasan Pengguna Aplikasi Myim3 Pada Situs Google Play Menggunakan Support Vector Machine,” *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 4, p. 1020, Aug. 2022, doi: 10.30865/jurikom.v9i4.4673.



- [11] R. A. Saputra, D. P. Ray, and F. Irwiensyah, “KLIK: Kajian Ilmiah Informatika dan Komputer Analisis Sentimen Aplikasi Tokocrypto Berdasarkan Ulasan Pada Google Play Store Menggunakan Metode Naïve Bayes,” *Media Online*, vol. 4, no. 4, 2024, doi: 10.30865/klik.v4i4.1707.
- [12] M. Y. Siregar, A. Davy Wiranata, and R. A. Saputra, “KLIK: Kajian Ilmiah Informatika dan Komputer Analisis Sentimen Pada Ulasan Pengguna Aplikasi Streaming Vidio Menggunakan Metode Naïve Bayes,” *Media Online*, vol. 4, no. 5, pp. 2419–2429, 2024, doi: 10.30865/klik.v4i5.1787.
- [13] R. S. Amardita, A. Adiwijaya, and M. D. Purbolaksono, “Analisis Sentimen terhadap Ulasan Paris Van Java Resort Lifestyle Place di Kota Bandung Menggunakan Algoritma KNN,” *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 1, p. 62, Feb. 2022, doi: 10.30865/jurikom.v9i1.3793.
- [14] E. Hasibuan and E. A. Heriyanto, “Analisis Sentimen Pada Ulasan Aplikasi Amazon Shopping Di Google Play Store Menggunakan Naive Bayes Classifier,” *Jurnal Teknik dan Science*, vol. 1, no. 3, pp. 13–24, Oct. 2022, doi: 10.56127/jts.v1i3.434.
- [15] N. Q. Rizkina and F. N. Hasan, “Analisis Sentimen Komentar Netizen Terhadap Pembubaran Konser NCT 127 Menggunakan Metode Naive Bayes,” *Journal of Information System Research (JOSH)*, vol. 4, no. 4, pp. 1136–1144, Jul. 2023, doi: 10.47065/josh.v4i4.3803.
- [16] Gilbert, Syariful Alam, and M. Imam Sulisty, “Analisis Sentimen Berdasarkan Ulasan Pengguna Aplikasi My Pertamina Pada Google Playstore Menggunakan Metode Naïve Bayes,” *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 2, no. 3, pp. 100–108, Aug. 2023, doi: 10.55123/storage.v2i3.2333.
- [17] H. Zhafran Muflih, A. Rizki Abdillah, and F. Noor Hasan, “Analisis Sentimen Ulasan Pengguna Aplikasi Ajaib Menggunakan Metode Naïve Bayes,” *Media Online*, vol. 4, no. 3, pp. 1613–1621, 2023, doi: 10.30865/klik.v4i3.1303.
- [18] K. Kevin, M. Enjeli, and A. Wijaya, “Analisis Sentimen Penggunaan Aplikasi Kinemaster Menggunakan Metode Naive Bayes,” *Jurnal Ilmiah Computer Science*, vol. 2, no. 2, pp. 89–98, Jan. 2024, doi: 10.58602/jics.v2i2.24.
- [19] R. Rizaldi and R. Aryanti, “Analisis Sentimen Pengguna Terhadap Aplikasi Indodana Di Google Play Store Menggunakan Metode Naive Bayes Classifier,” *Journal of Informatics Management and Information Technology*, vol. 3, no. 4, pp. 98–105, 2024, doi: 10.47065/jimat.v4i3.400.
- [20] M. Irfan and E. Erizal, “Perbandingan Algoritma Naïve Bayes dengan K-Nearest Neighbor Untuk Analisis Sentimen Aplikasi InDrive di Playstore,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 3, p. 1535, Jul. 2024, doi: 10.30865/mib.v8i3.7780.
- [21] R. Wahyudi and G. Kusumawardana, “Analisis Sentimen pada Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine,” *Jurnal Informatika*, vol. 8, no. 2, pp. 200–207, Sep. 2021, doi: 10.31294/ji.v8i2.9681.
- [22] A. Rhamadanti, A. Rifa’i, F. Dikananda, and K. Anam, “ANALISIS SENTIMEN PADA ULASAN ACCESS BY KERETA API INDONESIA DENGAN K-NEAREST NEIGHBOR,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 1, pp. 2830–7062, 2024, doi: 10.23960/jitet.v12i1.3691.
- [23] R. Maulana, A. Voutama, and T. Ridwan, “Analisis Sentimen Ulasan Aplikasi My Pertamina pada Google Play Store menggunakan Algoritma NBC,” *Jurnal Teknologi Terpadu*, vol. 9, no. 1, pp. 42–48, Jul. 2023, doi: 10.54914/jt.v9i1.609.
- [24] H. C. Husada and A. S. Paramita, “Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM),” *Teknika*, vol. 10, no. 1, pp. 18–26, Feb. 2021, doi: 10.34148/teknika.v10i1.311.
- [25] N. P. Husain, S. Sukirman, and S. SAJIAH, “Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine,” *Journal of System and Computer Engineering (JSCE)*, vol. 5, no. 1, pp. 91–102, Jan. 2024, doi: 10.61628/jsce.v5i1.1105.
- [26] F. N. Rahmaulidyah, M. N. Hayati, and R. Goejantoro, “Perbandingan Metode Klasifikasi Naive Bayes dan K-Nearest Neighbor pada Data Status Pembayaran Pajak Pertambahan Nilai di Kantor Pelayanan Pajak Pratama Samarinda Ulu,” *EKSPONENSIAL*, vol. 12, no. 2, p. 161, Dec. 2021, doi: 10.30872/eksponsional.v12i2.809.
- [27] E. W. Sholeha, S. Yunita, R. Hammad, V. C. Hardita, and K. Kaharuddin, “Analisis Sentimen Pada Agen Perjalanan Online Menggunakan Naïve Bayes dan K-Nearest Neighbor,” *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 3, no. 4, pp. 203–208, Jan. 2022, doi: 10.35746/jtim.v3i4.178.
- [28] I. P. Rahayu, A. Fauzi, and J. Indra, “Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Dan Support Vector Machine,” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 2, p. 296, Dec. 2022, doi: 10.30865/json.v4i2.5381.
- [29] M. N. Hidayat and R. Pramudita, “Analisis Sentimen Terhadap Pembelajaran Secara Daring Pasca Pandemi Covid-19 Menggunakan Metode IndoBERT,” *INFORMATION MANAGEMENT FOR EDUCATORS AND PROFESSIONALS: Journal of Information Management*, vol. 8, no. 2, p. 161, Jan. 2024, doi: 10.51211/imbi.v8i2.2719.
- [30] Irvandi, B. Irawan, and O. Nurdiawan, “Naive Bayes Dan Wordcloud Untuk Analisis Sentimen Wisata Halal Pulau Lombok,” *INFOTECH journal*, vol. 9, no. 1, pp. 236–242, May 2023, doi: 10.31949/infotech.v9i1.5322.