

Analisis Sentimen Pada Ulasan Pengguna Aplikasi Streaming Vidio Menggunakan Metode Naïve Bayes

Muhammad Yusuf Siregar, Ade Davy Wiranata*, Rizki Adi Saputra

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

Email: ¹siregaryusuf123@gmail.com, ^{2,*}adedavy@uhamka.ac.id, ³belajar7706@gmail.com

Email Penulis Korespondensi: adedavy@uhamka.ac.id

Abstrak—Perkembangan inovasi teknologi informasi positif bagi industri media dan hiburan, terutama dalam penggunaan layanan streaming online yang signifikan. Aplikasi streaming Vidio menjadi salah satu yang populer, menyediakan fitur menarik seperti menonton, mengunggah, dan berbagi video dalam format mobile untuk smartphone, tablet, dan smart TV. Kelebihan Vidio adalah kemampuannya menampilkan siaran televisi langsung dari berbagai saluran di Indonesia. Meskipun aplikasi Vidio menawarkan fitur-fitur yang menarik dan kemudahan dalam penggunaannya, tidak semua pengguna dijamin akan merasa puas. Setiap aplikasi memiliki keunggulan dan kekurangan masing-masing, sehingga pengalaman dan pandangan pengguna bisa bervariasi. Hal ini terlihat dari ulasan pengguna yang tersedia di Google Play Store. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna terhadap aplikasi Vidio menggunakan algoritma Naïve Bayes. Data yang digunakan dalam penelitian ini adalah 1500 ulasan yang dikumpulkan pada bulan Februari 2024. Setelah melalui tahap pembersihan data, jumlahnya berkurang menjadi 1475 data ulasan. Dari dataset yang tersedia, ada 165 ulasan positif dan 1310 ulasan negatif. Proses penelitian meliputi pengambilan data dengan teknik web scraping, pelabelan data, preprocessing, serta pembobotan TF-IDF sebelum tahap klasifikasi. Untuk klasifikasi, digunakan teknik cross validation dan metode Naïve Bayes. Hasil penelitian menunjukkan bahwa akurasi yang dicapai oleh algoritma Naïve Bayes adalah sebesar 80.28%, presisi sebesar 24.18%, dan recall sebesar 35.76%.

Kata Kunci: Naive Bayes; Analisis Sentimen; Aplikasi; Vidio; Streaming

Abstract—The development of information technology innovation has a positive impact on the media and entertainment industry, especially in the significant use of online streaming services. Vidio is one of the most popular streaming applications, providing interesting features such as watching, uploading, and sharing videos in mobile format for smartphones, tablets, and smart TVs. Vidio's strength is its ability to show live television from various channels in Indonesia. Although the Vidio app offers interesting features and ease of use, not all users are guaranteed to be satisfied. Each app has its own advantages and disadvantages, so users' experiences and views may vary. This can be seen from the user reviews available on the Google Play Store. This research aims to analyze the sentiment of user reviews of the Vidio application using the Naïve Bayes algorithm. The data used in this research are 1500 reviews collected in February 2024. After going through the data cleaning stage, the number was reduced to 1475 review data. From the available dataset, there are 165 positive reviews and 1310 negative reviews. The research process includes data retrieval using web scraping technique, data labeling, preprocessing, and TF-IDF weighting before the classification stage. For classification, cross validation technique and Naïve Bayes method were used. The results showed that the accuracy achieved by the Naïve Bayes algorithm was 80.28%, precision 24.18%, and recall 35.76%.

Keywords: Naive Bayes; Sentiment Analysis; Apps; Vidio; Streaming

1. PENDAHULUAN

Perkembangan inovasi dalam teknologi informasi memberikan dampak positif pada industri media dan hiburan dengan memperbesar penggunaan layanan *streaming* dari menonton televisi menjadi beralih ke layanan *streaming online*. Kegiatan menonton film melalui platform *streaming online* masih menjadi pilihan utama masyarakat saat merasa jenuh, yang kini menjadi kebiasaan yang tak terpisahkan dari gaya hidup modern. Layanan *streaming online* atau langganan serial digital menawarkan berbagai genre, bahasa, dan produksi dari negara-negara di seluruh dunia. Berbagai pilihan film dan serial yang disediakan tidak hanya menghibur, tetapi juga menarik dan seru, sehingga membuat penonton merasa ketagihan dan ingin menonton lagi[1]. Salah satu dari banyak *platform streaming online* adalah aplikasi *streaming* Vidio[2]. Aplikasi Vidio di platform ini menyajikan layanan *streaming* film untuk pengalaman hiburan dari Indonesia dan luar negeri, serta menawarkan pertandingan olahraga langsung[2]. Platform lokal yang populer di Indonesia ini merupakan hasil dari upaya PT Vidio Dot Com, yang saat ini dimiliki oleh Emtek[3]. Vidio telah berhasil mengalahkan pesaing asing seperti Netflix dan Disney+, menjadi platform *streaming* dengan pertumbuhan tercepat di Indonesia dan memiliki jumlah pengguna aktif terbesar, mencapai sekitar 60 juta penonton setiap bulannya[3].

Aplikasi Vidio bisa diunduh secara gratis melalui *Google Play Store*[4]. Aplikasi ini menampilkan antarmuka yang sederhana dan dilengkapi dengan berbagai fitur menarik yang mudah digunakan, serta memudahkan pengguna untuk berbagi konten dengan teman-teman mereka[5]. Aplikasi *streaming* Vidio ini menawarkan berbagai fitur menarik, seperti kemampuan untuk menonton, mengunggah, dan berbagi video. Vidio tersedia dalam format mobile yang dapat diunduh di *smartphone*, tablet, dan *smart TV*. Salah satu keistimewaan Vidio adalah kemampuannya untuk menonton siaran televisi secara langsung dari berbagai saluran televisi di Indonesia, menjadikannya berbeda dari platform streaming lainnya[4].

Walaupun Vidio menawarkan beragam fitur menarik untuk mempermudah penggunaannya, tidak bisa dipastikan bahwa semua pengguna akan merasa puas. Seperti halnya dengan setiap aplikasi lain, Vidio memiliki kelebihan dan kelemahan tersendiri, sehingga pengalaman dan pandangan pengguna bisa berbeda-beda. Ini terlihat dari berbagai ulasan yang diberikan oleh pengguna di *Google Play Store*[6]. *Google Play Store* sendiri merupakan platform untuk mengunduh berbagai aplikasi dan konten digital lainnya bagi pengguna smartphone *Android*, mulai dari *e-book*, *film*, *game*, dan lainnya, yang bisa diakses baik secara gratis maupun dengan membayar[7]. Salah satu fitur penting di *Google Play Store* adalah rating dan ulasan. Fitur ini memiliki potensi untuk memengaruhi keputusan calon pengguna, karena kebanyakan pengguna cenderung melihat ulasan sebagai parameter untuk menilai seberapa baik suatu aplikasi. Ulasan juga dapat menjadi pertimbangan bagi pengguna sebelum mereka memutuskan untuk mengunduh aplikasi tersebut[8]. Dari ulasan tersebut, perusahaan perlu meningkatkan mutu produk yang telah diluncurkannya. Untuk mendapatkan pemahaman mengenai respons dan tanggapan dari pengguna aplikasi *streaming* Vidio, digunakan analisis sentimen[9]. Data dari ulasan pengguna di *Google Play Store* menjadi sumber utama dalam analisis ini [6].

Analisis sentimen, yang juga dikenal sebagai "opini mining" adalah suatu proses untuk mengidentifikasi pandangan atau opini seseorang yang diungkapkan dalam bentuk data teks, yang kemudian dapat diklasifikasikan sebagai sentimen positif atau negatif[10]. Melalui analisis sentimen, akan diperoleh gambaran menyeluruh tentang bagaimana pendapat pengguna mengenai pengalaman mereka menggunakan aplikasi Vidio[6]. Dalam studi ini, digunakan metode algoritma *Naïve Bayes*. Metode ini menerapkan pendekatan probabilitas dan statistik dengan memanfaatkan teorema *Bayes* yang disederhanakan, dimana diasumsikan bahwa setiap atribut memiliki independensi yang terpisah. Dengan kata lain, algoritma ini berasumsi bahwa fitur-fitur dari kelas tertentu dapat berdiri sendiri atau memiliki ketergantungan yang minimal satu sama lain[11]. *Naïve Bayes* digunakan untuk mengelompokkan ulasan pengguna ke dalam kategori sentimen positif dan negatif karena kemudahannya, kecepatannya, serta tingkat akurasi yang tinggi[12].

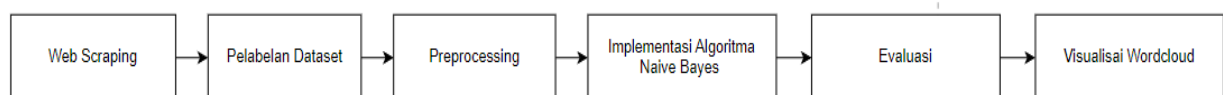
Dalam studi sebelumnya pada tahun 2023, Hilmy Zhafran Muflih mengulas analisis sentimen terhadap Ulasan Pengguna Aplikasi Ajaib dengan menggunakan Metode *Naïve Bayes*, ditemukan bahwa dari 500 data yang dianalisis, tingkat akurasinya mencapai 76,67%.[13]. Berdasarkan penelitian yang dilakukan oleh Tia Anggita Sari, Estu Sinduningrum, dan Firman Noor Hasan dalam penelitian berjudul "Analisis Sentimen Ulasan Pelanggan Pada Aplikasi Fore Coffee Menggunakan Metode *Naïve Bayes*", hasil akurasinya mencapai 74,28%. Selain itu, presisi yang tercatat adalah sebesar 24,46%, dan *recall* sebesar 45,33%[14]. Tingkat akurasi sebesar 87 dicatat dalam penelitian sebelumnya oleh Mahardika Tania Nitami dan Hery Februariyanti, yang menggunakan algoritma *Naïve Bayes* untuk analisis sentimen ulasan ekspedisi J&T Express[15]. Studi analisis sentimen, yang juga memanfaatkan algoritma *Naïve Bayes*, dilakukan oleh Sendi Alpin Rizaldi dan timnya terhadap aplikasi Jamsostek Mobile pada tahun 2023. Hasil penelitian tersebut menunjukkan bahwa ulasan pengguna aplikasi cenderung negatif, dengan tingkat akurasi mencapai 95%[16]. Penelitian yang dilakukan oleh Mujaddid Izzul Fikri dan timnya mencakup perbandingan antara metode *Naïve Bayes* dan metode lain, seperti *Support Vector Machine* (SVM), dalam analisis sentimen dengan menggunakan data tweet yang terkait dengan Universitas Muhammadiyah Malang. Hasil penelitian menunjukkan akurasi sebesar 73.65% untuk metode *Naïve Bayes* dan 70.20% untuk metode *support vector machine*[17].

Untuk memastikan bahwa aplikasi layanan *streaming* Vidio dapat memberikan manfaat optimal dan menciptakan dampak positif bagi pengguna, diperlukan penelitian lebih lanjut. Sebagai contoh, penelitian analisis sentimen terhadap aplikasi *streaming* Mola di *Google Play Store* yang dilakukan oleh Novi Lestari dan tim menggunakan Algoritma *Naïve Bayes* pada tahun 2023 menunjukkan bahwa terdapat lebih banyak sentimen negatif, dengan tingkat akurasi mencapai 64%, presisi sebesar 58%, dan *recall* sebesar 89%[18].

Dari uraian masalah sebelumnya, fokus penelitian ini adalah pada analisis sentimen menggunakan metode *Naïve Bayes*. Pemilihan metode ini didasarkan pada hasil penelitian sebelumnya yang telah terbukti memiliki tingkat akurasi yang tinggi, bahkan melebihi algoritma lain seperti *support vector machine*, sebagaimana yang didokumentasikan dalam penelitian sebelumnya oleh Mujaddid Izzul Fikri. Penelitian ini memperhitungkan akurasi, presisi, dan *recall* yang tercantum dalam *confusion matrix* untuk mengevaluasi keakuratan terhadap metode *Naïve Bayes*. Tujuan utama penelitian ini adalah untuk memberikan gambaran pandangan pengguna terhadap penggunaan aplikasi *streaming* Vidio. Oleh karena itu, memahami pendapat dan penilaian pengguna terhadap aplikasi tersebut penting, karena informasi tersebut dapat membimbing pengembang dalam meningkatkan kualitas dan manfaatnya bagi pengguna aplikasi.

2. METODOLOGI PENELITIAN

Tahapan yang dilakukan dalam penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna terhadap aplikasi *streaming* Vidio di kolom ulasan *Google Play Store* menggunakan metode *Naïve Bayes*, sebagaimana ditunjukkan pada Gambar 1.



Gambar 1. Tampilan Dari Alur Penelitian

Pada langkah-langkah penelitian yang tergambar dalam gambar 1, dimulai dengan pengumpulan data ulasan pengguna aplikasi streaming Vidio di *Google Play Store* menggunakan teknik *scrapping* dengan bantuan *Google Colab*. Selanjutnya, data akan dilabeli secara manual sebelum menjalani serangkaian tahap *preprocessing*, termasuk *cleaning*, *tokenize*, *transform cases*, *filter stopwords*, dan *filter token by length*, yang kemudian diikuti oleh pembobotan TF-IDF. Setelah proses ini selesai, data yang telah dipersiapkan akan diimplementasikan dengan menggunakan metode *Naïve Bayes*. Tahap terakhir dari penelitian ini adalah evaluasi menggunakan *confusion matrix* untuk mengukur akurasi, presisi, dan recall, serta melakukan visualisasi *wordcloud*.

2.1 Web Scraping

Web scraping merujuk pada teknik otomatis untuk mengambil data atau informasi dari situs web. Tujuan dari *web scraping* adalah mengekstraksi informasi dari berbagai situs web yang bersifat tidak terstruktur, kemudian mentransformasikannya menjadi format yang lebih terstruktur dan teratur[19]. Data yang diperoleh melalui proses *web scraping* kemudian disimpan dan diubah formatnya menjadi file CSV.

2.2 Pelabelan Data

Data yang diperoleh dari teknik web scraping kemudian dianalisis dengan memberikan label pada setiap ulasan, dan diproses secara manual dalam format *Comma Separated Values (CSV)* menggunakan *Microsoft Excel*. Proses penambahan label pada data ini penting untuk mengidentifikasi kategori dari setiap ulasan dalam dokumen, apakah komentar atau ulasan yang didapat itu termasuk dalam kategori positif atau negatif. Kategori positif mencakup pujian, saran, dan ekspresi positif lainnya yang menunjukkan kepuasan pengguna terhadap aplikasi. Di sisi lain, kategori negatif meliputi keluhan, protes, keberatan, dan perasaan negatif seperti ketidakpuasan, dan kekecewaan[20].

2.3 Preprocessing

Langkah *preprocessing* melibatkan persiapan dataset dengan membersihkan dan menyusun kontennya dari dataset mentah menjadi dataset yang terorganisir dan bersih. Proses *preprocessing* ini melibatkan penggunaan dataset yang telah diberi label sebelumnya[21]. Langkah ini ditempuh dengan maksud supaya dataset bisa diolah secara efektif selama proses klasifikasi, dengan mempersiapkan teks yang akan diolah selanjutnya melalui transformasi data dari bentuk yang belum atau tidak teratur menjadi data yang teratur[22]. Terdapat beberapa tahapan dalam proses *preprocessing* yang terdiri dari beberapa tahap, diantaranya:

a. Cleaning

Pada tahap ini, data dimurnikan dengan menghapus informasi yang tidak relevan, seperti *URL*, tagar, tanda baca, nama pengguna, angka, dan karakter yang tidak perlu[23]. Setelah proses pembersihan, dataset akan disisakan sesuai dengan kebutuhan yang diinginkan[12].

b. Transform cases

Adalah proses mengubah karakter alfabet, yang telah melewati tahap pembersihan, menjadi huruf kecil[24].

c. Tokenize

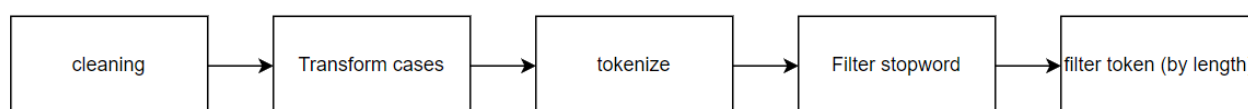
Tokenisasi, atau *tokenize*, adalah proses yang bertujuan untuk memecah rangkaian kata dalam suatu kalimat, paragraf, atau halaman menjadi segmen-segmen kata yang lebih kecil[25]. Proses pemisahan kalimat menjadi kata dilakukan dengan cara memotong kalimat berdasarkan spasi[26].

d. Filter Stopwords

Penggunaan *filter stopwords* bertujuan untuk menghilangkan kata penghubung atau imbuhan seperti 'dan', 'yang', atau 'ke', dan hanya menyisakan kata-kata penting yang memiliki makna[27].

e. Filter token (by length)

Pada tahap ini, kata-kata dengan jumlah karakter tertentu akan dihilangkan[28]. Peneliti memberlakukan batasan panjang huruf, yang mengharuskan kata harus memiliki setidaknya empat karakter huruf dan tidak boleh melebihi 25 huruf per kata.



Gambar 2. Tahapan *Preprocessing*

Gambar 2 menunjukkan tahap-tahap dalam *preprocessing* data, yang dimulai dari proses *cleaning*, *transform cases*, *tokenize*, *filter stopwords*, hingga *filter token by length*.

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

Penelitian ini menggunakan teknik pemrosesan bahasa alami yang disebut *Frequency-Inverse Document Frequency (TF-IDF)* untuk menentukan tingkat kepentingan sebuah kata dalam dokumen atau kumpulan dokumen yang disebut korpus.

Metode ini menggabungkan dua elemen utama yaitu *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF). *Term Frequency* mengukur seberapa sering suatu kata muncul dalam dokumen dengan membandingkan jumlah kemunculan kata tersebut dengan total jumlah kata dalam dokumen[29]. Sementara itu, IDF menilai seberapa jarang atau unik suatu kata di seluruh korpus dengan membandingkan logaritma total jumlah dokumen dengan jumlah dokumen yang mengandung kata tersebut[30].

2.5 Implementasi Algoritma Naive Bayes

Algoritma *Naive Bayes* adalah metode klasifikasi yang digunakan untuk meramalkan probabilitas dan statistik. Metode ini memproyeksikan probabilitas bahwa data tertentu termasuk dalam suatu kelas berdasarkan perhitungan probabilitasnya. Algoritma ini didasarkan pada *Teorema Bayes*, yang ditemukan oleh *Thomas Bayes*[31]. *Naive Bayes* sangat bermanfaat dalam klasifikasi teks, terutama dalam analisis sentimen di mana tujuannya adalah mengenali apakah dokumen atau teks memiliki sifat positif atau negatif berdasarkan kata-kata yang terkandung di dalamnya[32]. Berikut adalah rumus perhitungan yang digunakan dalam metode *Naive Bayes*.

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

Dalam rumus ini variabel X mewakili data dokumen dengan kelas yang belum diketahui, sementara H adalah hipotesis yang menyatakan kelas spesifik dari data X. $P(H|X)$ adalah probabilitas hipotesis H berdasarkan kondisi X, yang juga dikenal sebagai posteriori probability. Selain itu, P(H) menunjukkan probabilitas dari hipotesis H sebelum mempertimbangkan data X, atau dikenal sebagai prior probability. Probabilitas dari data X jika hipotesis H benar dinyatakan sebagai $P(X|H)$, sedangkan P(X) adalah probabilitas data X secara umum [31].

2.6 Evaluasi

Langkah selanjutnya adalah menggunakan confusion matrix untuk menilai kinerja metode klasifikasi sebagai bagian dari evaluasi[33]. Ini adalah alat yang umum digunakan untuk menghitung tingkat akurasi. Dalam pengujian keakuratan hasil pencarian, akan dinilai nilai recall, presisi, akurasi, dan *error rate*[33]. Rumus untuk menghitung *confusion matrix* adalah sebagai berikut :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

Penjelasan dari rumus tersebut mencakup *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), dan *True Negative* (TN). TP menunjukkan jumlah data yang benar-benar positif dan diprediksi dengan benar sebagai positif, sedangkan FP adalah jumlah data yang sebenarnya negatif tetapi salah diprediksi sebagai positif. Sebaliknya, FN mengacu pada jumlah data yang sebenarnya positif tetapi diprediksi dengan salah sebagai negatif, sedangkan TN menunjukkan jumlah data yang benar-benar negatif dan diprediksi dengan benar sebagai negatif[33].

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

Gambar 3. Tampilan *confusion matrix*

Pada Gambar 3 terdapat *class prediction* dari *confusion matrix* dalam analisis ini, digunakan variabel seperti True TP, FP, FN, dan TN.

2.7 Visualisasi Wordcloud

Dalam studi ini, Peneliti memanfaatkan Visualisasi yang dipresentasikan dalam format *wordcloud*. *Wordcloud* adalah metode visualisasi data yang mengilustrasikan informasi dengan memprioritaskan kata-kata berdasarkan frekuensi kemunculannya dalam teks tertulis. Pemanfaatan *wordcloud* dalam analisis bahasa membantu peneliti dengan memberikan gambaran instan tentang inti dari teks[34].

3. HASIL DAN PEMBAHASAN

3.1 Scraping Data

Pada tahap ini, peneliti menghimpun data ulasan dari *Google Play Store* pada tanggal 19 Februari 2024. Penggunaan metode *Web Scraping* dilakukan dengan bahasa pemrograman *Python*. Peneliti harus memilih jenis data yang akan dipilih seperti rating, nama, jumlah data dan terutama ulasan pada aplikasi.



Gambar 4. Proses Pengambilan Data Ulasan

Gambar 4 memperlihatkan Proses penghimpunan data ulasan pada penelitian ini dilakukan melalui teknik *web scraping* dengan mengakses *Google Play Store*, kemudian menuju aplikasi *Vidio* untuk mendapatkan link yang akan digunakan selanjutnya, mengakses *Google Colab*, dan menempelkan link tersebut untuk melakukan *scraping data* yang kemudian akan disimpan dalam format *CSV* dengan bantuan *Google Colab*. Data yang diambil hanya mencakup ulasan yang berkaitan dengan aplikasi *streaming Vidio*. Hasil dari proses pengumpulan data melalui teknik *Scraping* menggunakan *Google Colab* adalah sebagai berikut:

```
#scrape jumlah ulasan yang diinginkan
from google_play_scraper import Sort, reviews

result, continuation_token = reviews(
    'com.vidio.android',
    lang='id', #Indonesia
    country='id', #bahasa Indonesia
    sort=Sort.MOST_RELEVANT, # #relevan
    count=1500, # jumlah ulasan
    filter_score_with=None # # kemudian di filter_score
)
```

Gambar 5. Proses Scraping Data Ulasan Pada Aplikasi Vidio

Pada Gambar 5 di atas memperlihatkan langkah-langkah dalam mengambil data ulasan aplikasi *Vidio* menggunakan *Google Colab* dengan menautkan link 'com.vidio.android'. Hasil dari langkah tersebut mencakup 1500 ulasan komentar pada aplikasi *Vidio* dari *Google play store*, dengan fokus pada pemilihan komentar berbahasa Indonesia, berasal dari negara Indonesia, dan menekankan pada komentar yang paling relevan. Dataset yang sudah diambil akan diubah ke format *CSV*. Setelah itu, pembersihan manual dilakukan untuk menghapus komentar yang tidak relevan, mengurangi jumlah data dari 1500 menjadi 1477 ulasan.

3.2 Pelabelan Data

Selama tahap pelabelan data, peneliti memberikan label pada setiap komentar yang diperoleh melalui *web scraping*. Pelabelan tersebut terdiri dari dua jenis yaitu label positif dan label negatif. Dalam menganalisis sentimen di aplikasi *Vidio*, ditemukan beberapa komentar yang menghasilkan sentimen negatif dan positif.

Tabel 1. Hasil Pelabelan Data Ulasan Aplikasi Vidio

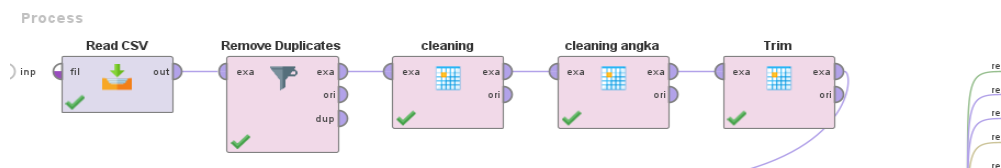
Label	Ulasan
Negatif	Terlalu banyak iklan dan iklanya lama
Negatif	sering gangguan, sering buffering parah. bikin gak nyaman nonton
Positif	Saya pikir aplikasinya udah bagus
Positif	Aplikasi nya cukup bagus dan nyaman di pakai

Tabel 1 menampilkan hasil pelabelan data. Pada bagian komentar negatif, terdapat beberapa hal yang tidak disukai pengguna aplikasi, seperti kesulitan dalam menonton, keluhan tentang banyaknya iklan yang muncul, dan masalah koneksi yang sering terjadi saat menggunakan aplikasi *Vidio*. Selain itu, pelanggan yang berlangganan merasa frustrasi karena iklan masih muncul, dianggap tidak bermanfaat, bahkan merugikan. Di sisi lain, pada kategori komentar positif, beberapa pengguna menyatakan kepuasan mereka. Mereka menemukan kemudahan untuk menemukan konten yang diinginkan dan menikmati fitur-fitur yang berguna dalam aplikasi *Vidio*.

3.3 Preprocessing

Dalam tahap penelitian ini, dilakukan persiapan data dimana data yang terkumpul akan menjalani proses *cleaning* untuk menghilangkan data duplikat. Setelah itu, dataset ulasan aplikasi *Vidio* yang mengandung sentimen akan dilanjutkan ke tahap *preprocessing* dengan menggunakan alat dari *RapidMiner*. *Preprocessing* merupakan langkah penting untuk mengubah kalimat yang tak terstruktur menjadi kalimat yang terstruktur, memungkinkan penggunaan data secara optimal.

Langkah-langkah dalam proses *preprocessing* dataset mencakup *tokenize*, *transform cases*, *filter stopwords*, dan *filter token by length*.



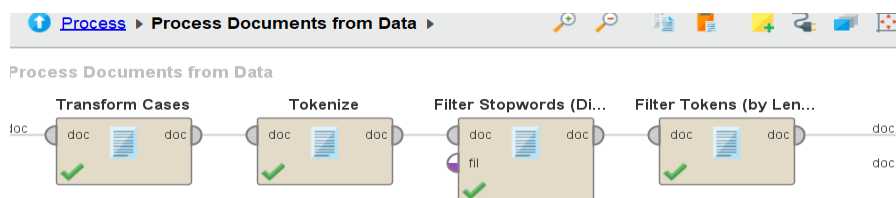
Gambar 6. Proses *Preprocessing Cleaning* data

Pada Gambar 6 kita dapat melihat langkah-langkah pembersihan data yang telah dilakukan dan disimpan dalam file *CSV*. Pembersihan data diperlukan karena hasil ulasan mengandung simbol-simbol seperti `[;:'"!@#$$%*+=]`. Kemudian dilakukan juga pembersihan angka. Setelah itu, digunakan operator *Remove Duplicates* untuk menghapus data duplikat yang mungkin muncul selama proses pengambilan data serta menambahkan operator *Trim* untuk menghilangkan spasi ganda dari data ulasan yang ada.

Tabel 2. Hasil *Cleaning* Data

Komentar sebelum <i>cleaning</i>	Komentar sesudah <i>cleaning</i>
Setelah update, kok ada bug yah, setelah menonton kurang dari 1 menit, keluar sendiri dari apk nya, dan patah" Saat menonton, tolong di perbaiki lagi	Setelah update kok ada bug yah setelah menonton kurang dari menit keluar sendiri dari apk nya dan patah Saat menonton tolong di perbaiki lagi
Tolong dong perbaiki lagi aplikasi nya jangan bikin penonton ga nyaman. Masa setiap nonton gangguan terus. Harus di ulang2 bikin emosi.	Tolong dong perbaiki lagi aplikasi nya jangan bikin penonton ga nyaman Masa setiap nonton gangguan terus Harus di ulang bikin emosi

Pada Tabel 2, kolom kiri menampilkan komentar sebelum dilakukan *cleaning*, sedangkan kolom kanan menunjukkan hasil komentar setelah proses *cleaning*, yaitu hasil dari pembersihan data. Pada teks ulasan sebelum pembersihan, terdapat simbol atau karakter seperti `‘./’!` dan angka yang kemudian dihapus, sehingga hanya kata-kata yang tersisa tanpa simbol atau karakter tersebut. Setelah dilakukan pembersihan, dari 1477 data ulasan, tersisa 1475 data. Dari jumlah tersebut, terdapat 1310 ulasan negatif dan 165 ulasan positif.



Gambar 7. Operator Pada *Preprocessing* Data

Selain itu, pada gambar 7 yaitu penjelasan tentang *preprocessing* data dalam penelitian ini melibatkan penggunaan beberapa operator pemrosesan dokumen pada RapidMiner. Di antara operator-operator yang digunakan adalah *Tokenize*, *Transform Cases*, *Filter Stopwords*, dan *Filter Tokens by Length*.

a. *Transform Cases*

Ini adalah langkah dalam proses *Transform cases*, di mana karakter kata yang awalnya berhuruf besar setelah melalui tahap pembersihan akan diubah menjadi huruf kecil. Tabel 3 di bawah akan menampilkan beberapa komentar sebelum dan setelah melewati proses *transform cases*.

Tabel 3. Hasil *Transform Cases*

Komentar sebelum <i>transform cases</i>	Komentar sesudah <i>transform cases</i>
Setelah update kok ada bug yah setelah menonton kurang dari menit keluar sendiri dari apk nya dan patah Saat menonton tolong di perbaiki lagi	setelah update kok ada bug yah setelah menonton kurang dari menit keluar sendiri dari apk nya dan patah saat menonton tolong di perbaiki lagi
Tolong dong perbaiki lagi aplikasi nya jangan bikin penonton ga nyaman. Masa setiap nonton gangguan terus. Harus di ulang bikin emosi	tolong dong perbaiki lagi aplikasi nya jangan bikin penonton ga nyaman masa setiap nonton gangguan terus harus di ulang bikin emosi

Pada Tabel 3, sebelum dilakukan transformasi huruf atau *transform cases*, terdapat huruf besar seperti "S", "T", "H", dan lain-lain. Setelah proses transformasi, huruf-huruf besar tersebut berubah menjadi huruf kecil, sebagaimana yang terlihat pada tampilan tabel sesudah *transform cases*.

b. *Tokenize*

Pada tahap ini, tujuan dari proses *Tokenize* atau tokenisasi adalah untuk memisahkan urutan kata dalam kalimat, paragraf, atau halaman menjadi potongan-potongan kata yang terpisah. Proses pemisahan kalimat menjadi kata dilakukan

dengan cara memotong kalimat berdasarkan spasi. Tabel 4 memperlihatkan komentar sebelum dilakukan *tokenize* dan sesudah dilakukan *tokenize*.

Tabel 4. Hasil *Tokenize*

Komentar sebelum <i>tokenize</i>	Komentar sesudah <i>tokenize</i>
setelah update kok ada bug yah setelah menonton kurang dari menit keluar sendiri dari apk nya dan patah saat menonton tolong di perbaiki lagi	'setelah', 'update', 'kok', 'ada', 'bug', 'yah', 'setelah', 'menonton', 'kurang', 'dari', 'menit', 'keluar', 'sendiri', 'dari', 'apk', 'nya', 'dan', 'patah', 'saat', 'menonton', 'tolong', 'di', 'perbaiki', 'lagi'
tolong dong perbaiki lagi aplikasi nya jangan bikin penonton ga nyaman masa setiap nonton gangguan terus harus di ulang bikin emosi	'tolong', 'dong', 'perbaiki', 'lagi', 'aplikasi', 'nya', 'jangan', 'bikin', 'penonton', 'ga', 'nyaman', 'masa', 'setiap', 'nonton', 'gangguan', 'terus', 'harus', 'di', 'ulang', 'bikin', 'emosi'

Pada Tabel 4 terdapat proses *tokenize* yang bertujuan untuk memisahkan bagian-bagian kata dalam sebuah kalimat. Setiap ulasan diproses dengan cara memisahkan potongan-potongan kata menggunakan proses *tokenize* dengan operator Rapidminer.

c. *Filter Stopwords*

Setelah dilakukan *tokenize* maka selanjutnya akan dilakukan tahap *filter stopwords* yaitu suatu prosedur yang dilakukan untuk menghilangkan setiap kata penghubung dan istilah yang tidak perlu dari kumpulan data. Penggunaan operator *filter stopwords* pada alat RapidMiner memungkinkan proses eliminasi kata-kata tidak relevan berdasarkan kamus Bahasa Indonesia yang diunduh dari situs Kaggle.

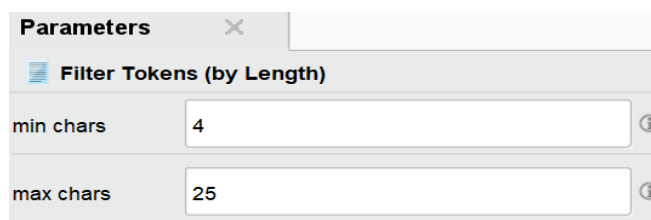
Tabel 5. Hasil *Filter stopwords*

Data sebelum <i>filter removal</i>	Data sesudah <i>filter removal</i>
'setelah', 'update', 'kok', 'ada', 'bug', 'yah', 'setelah', 'menonton', 'kurang', 'dari', 'menit', 'keluar', 'sendiri', 'dari', 'apk', 'nya', 'dan', 'patah', 'saat', 'menonton', 'tolong', 'di', 'perbaiki', 'lagi'	'update', 'bug', 'yah', 'menonton', 'menit', 'apk', 'nya', 'patah', 'menonton', 'tolong', 'perbaiki'
'tolong', 'dong', 'perbaiki', 'lagi', 'aplikasi', 'nya', 'jangan', 'bikin', 'penonton', 'ga', 'nyaman', 'masa', 'setiap', 'nonton', 'gangguan', 'terus', 'harus', 'di', 'ulang', 'bikin', 'emosi'	'tolong', 'perbaiki', 'aplikasi', 'nya', 'bikin', 'penonton', 'ga', 'nyaman', 'nonton', 'gangguan', 'ulang', 'bikin', 'emosi'

Kata-kata seperti "kok", "ada", "dong", dan lainnya masih terdapat pada Tabel 5 sebelum dilakukan penyaringan stopwords. Namun, setelah dilakukan tahap *filter stopwords*, kata-kata tersebut telah dihapus seperti yang terlihat pada tabel kanan sesudah *filter stopwords*.

d. *Filter token (by length)*

Langkah berikutnya akan memasuki tahap *filter token (by length)* berdasarkan panjangnya suatu huruf. Tujuan dari proses *Filter token by length* adalah untuk membatasi jumlah karakter dalam kalimat data yang sedang diproses dengan cara menyaring token berdasarkan panjangnya.



Gambar 8. Tampilan Operator Pembatasan karakter

Pada Gambar 8 menampilkan operator dari *filter token by length* saat proses menetapkan batasan huruf dengan minimum 4 dan maksimum 25 karakter.

Tabel 6. Hasil *filter token by Length*

Data sebelum <i>filter token by length</i>	Data sesudah <i>filter token by length</i>
'update', 'bug', 'yah', 'menonton', 'menit', 'apk', 'nya', 'patah', 'menonton', 'tolong', 'perbaiki'	'update', 'menonton', 'menit', 'patah', 'tolong', 'perbaiki'
'tolong', 'perbaiki', 'aplikasi', 'nya', 'bikin', 'penonton', 'ga', 'nyaman', 'nonton', 'gangguan', 'ulang', 'bikin', 'emosi'	'tolong', 'perbaiki', 'aplikasi', 'bikin', 'penonton', 'nyaman', 'nonton', 'gangguan', 'ulang', 'bikin', 'emosi'

Setelah proses *filter token by length* dilakukan pada Tabel 6, kata-kata dengan kurang dari 4 karakter seperti "yah", "nya", "apk", "bug", dan lain-lain telah dihilangkan dari kalimat tersebut.

3.4 Pembobotan Term Frequency - Inverse Document Frequency (TF-IDF)

Pada tahap ini, dilakukan pembobotan dalam proses TF-IDF dilakukan dengan menggunakan data hasil *preprocessing* yang telah disiapkan. Langkah pertama adalah pencarian *Term Frequency* (TF) untuk mengukur frekuensi kemunculan setiap kata dalam dokumen. TF memberikan indikasi seberapa sering suatu kata muncul dalam dokumen tertentu melalui rumus khusus. Setelah itu, dilakukan perhitungan *Inverse Document Frequency* (IDF).

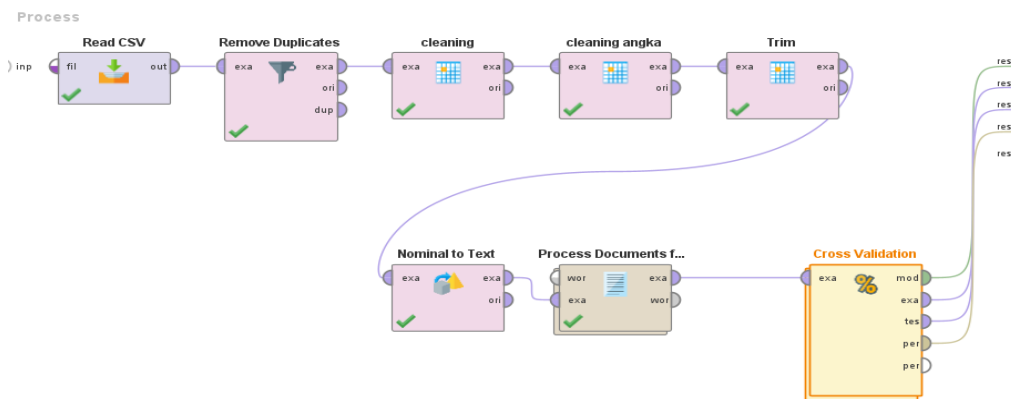
Tabel 7. Tampilan Dari Proses Pembobotan TF-IDF

Text	abis	aplikasi	bagus	iklan	buruk
abis update patah patah	0.457	0	0	0	0
banget iklan aplikasi nonton	0	0.406	0	0	0
bagus cuman sayang iklannya	0	0	0.270	0	0
kebanyakan iklan vidio iklan	0	0	0	0.609	0
buruk aplikasinya serakah iklan	0	0	0	0.071	0.192
bagus menghibur nonton vidio	0	0	0.297	0	0
bagus banget fiturnya udah bagus iklan menarik	0	0	0.276	0.118	0

Tabel 7 menampilkan hasil penilaian TF-IDF, dimana masing-masing kata diberi nilai bobot berdasarkan frekuensi kemunculannya dalam dokumen.

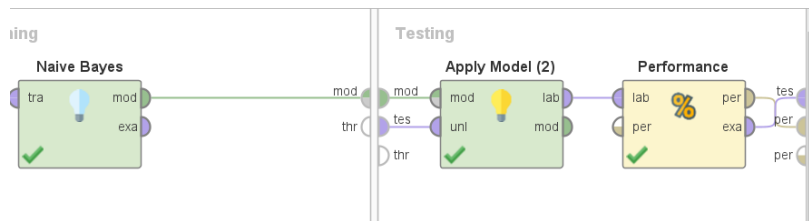
3.5 Algoritma Naïve Bayes

Setelah dataset berhasil disiapkan melalui proses preprocessing, peneliti melakukan pengujian *cross-validation* untuk mengevaluasi hasil klasifikasi dan akurasi dataset saat menerapkan algoritma *Naive Bayes*. Proses ini memanfaatkan operator *cross-validation*, dengan menerapkan teknik *k-fold cross-validation* dengan nilai k sebesar 10. Hal ini bertujuan untuk membagi data menjadi 10 subset yang setara.



Gambar 9. Proses Klasifikasi Data

Pada gambar 9, disajikan semua langkah dalam proses klasifikasi data menggunakan algoritma *Naive Bayes*. Semua operator terkoneksi dari Dataset berlabel hingga *Cross Validation*, dengan tujuan mengevaluasi kinerja algoritma tersebut.



Gambar 10. Tampilan Operator Cross Validation

Gambar 10 menampilkan operator *Cross Validation*, yang menggunakan algoritma *Naive Bayes* untuk klasifikasi data. Operator *Cross Validation* akan menampung beberapa operator lainnya, termasuk *Naive Bayes*, *Apply Model*, dan *Performance*. Fungsi dari *Apply Model* adalah menerima data input dan menghasilkan prediksi keluaran berdasarkan metode klasifikasi. Operator *Performance* digunakan untuk mengevaluasi kinerja algoritma *Naive Bayes* dalam melakukan klasifikasi data.

3.6 Evaluasi

Berdasarkan evaluasi pemodelan algoritma *Naïve Bayes* menggunakan alat RapidMiner, diperoleh tingkat akurasi sebesar 80,28%. *Confusion matrix* menunjukkan bahwa terdapat 59 *True Positive* (TP), 185 *False Positive* (FP), 106 *False Negative* (FN), dan 1125 *True Negative* (TN).

accuracy: 80.28% +/- 3.29% (micro average: 80.27%)			
	true Negative	true Positive	class precision
pred. Negative	1125	106	91.39%
pred. Positive	185	59	24.18%
class recall	85.88%	35.76%	

Gambar 11. Tampilan Hasil Evaluasi Naive Bayes

Gambar 11 merupakan visualisasi dari hasil evaluasi algoritma *Naïve Bayes* pada tahap klasifikasi data menggunakan perangkat RapidMiner. Visualisasi ini memberikan gambaran tentang kinerja algoritma dalam mengklasifikasikan data. Dibawah ini adalah hasil perhitungan untuk mengukur *accuracy*, *precision*, dan *recall* menggunakan rumus algoritma Naive Bayes dengan menggunakan *confusion matrix* dari hasil data:

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} = \frac{(59+1125)}{(59+1125+185+106)} = \frac{1184}{1475} = 0.8027 = 80.27\% \text{ atau } 80.28\%$$

$$Precision = \frac{TP}{(TP+FP)} = \frac{59}{(59+185)} = \frac{59}{244} = 0.2418 = 24.18\%$$

$$Recall = \frac{TP}{(TP+FN)} = \frac{59}{(59+106)} = \frac{59}{165} = 0.3576 = 35.76\%$$

Berdasarkan perhitungan dari *confusion matrix* di atas, hasil evaluasi klasifikasi data ulasan pengguna di aplikasi Vidio, diperoleh nilai akurasi sebesar 80,28%, *precision* sebesar 24,18%, dan *recall* sebesar 35,76%.

3.7 Visualisasi Wolrdcloud

Visualisasi kata pada tahap ini ditampilkan menggunakan *Wordcloud*, sebuah representasi visual dari teks di mana kata-kata yang sering muncul memiliki ukuran lebih besar, sementara kata-kata yang jarang muncul memiliki ukuran lebih kecil.



Gambar 12. Tampilan Visualisai Wordcloud

Gambar 12 menunjukkan representasi visual dalam bentuk *wordcloud* yang menyoroti kata-kata dengan frekuensi kemunculan antara 1 hingga 25 kata setelah data diurutkan secara menurun. *Wordcloud* ini mencerminkan bahwa kata 'nonton' mendominasi, dengan frekuensi kemunculan sebanyak 475 kali. Posisi kedua dipegang oleh kata 'aplikasi' yang muncul sebanyak 394 kali, sementara kata 'iklan' menempati posisi ketiga dengan kemunculan sebanyak 319 kali, kemudian kata 'bagus' muncul sebanyak 248 kali.

4. KESIMPULAN

Berdasarkan hasil dari penelitian diatas menunjukkan bahwa pada bulan Februari 2024, sejumlah 1500 data ulasan aplikasi Vidio dari *Google Play Store* berhasil terkumpul melalui teknik *web scraping* dengan menggunakan *Google*

Colab. Kemudian data yang diambil adalah dalam bentuk komentar berbahasa Indonesia yang berasal dari Indonesia, difokuskan pada yang paling relevan, dan kemudian disimpan dalam format file CSV. langkah berikutnya adalah pelabelan data yang dimana dari 1500 data ulasan tersebut setelah dilakukan proses pembersihan tersisa 1477 data. Setelah dilakukan pembersihan tambahan menggunakan alat RapidMiner, dari 1477 data ulasan, tersisa 1475 data. Dari jumlah tersebut, terdapat 165 ulasan positif dan 1310 ulasan negatif. Keluhan utama dalam ulasan negatif adalah mengenai iklan yang mengganggu, sementara pelanggan merasa senang dengan aplikasi yang memudahkan mereka menonton acara favorit dan memiliki fitur yang baik dalam ulasan positif. Setelah menyelesaikan serangkaian tahapan pengumpulan data, langkah selanjutnya adalah melakukan *preprocessing*. Ini melibatkan proses *cleaning* data untuk menghilangkan duplikat, menghapus simbol dan angka, kemudian dilanjutkan dengan *tokenize*, *transform cases*, *filter stopwords*, dan *filter token by length* untuk menetapkan batas minimum dan maksimum karakter. Sebelum tahap klasifikasi dimulai, langkah pembobotan kata diterapkan menggunakan teknik *TF-IDF*. Proses klasifikasi kemudian diterapkan melalui *cross validation*, terutama dengan menerapkan pendekatan *10-fold cross validation* dan memanfaatkan algoritma *Naïve Bayes*. Hasil klasifikasi menunjukkan tingkat akurasi sebesar 80.28%, presisi sebesar 24.18%, dan recall sebesar 35.76%. Untuk saran terhadap pengembangan penelitian ini, disarankan membandingkan klasifikasi dengan algoritma lainnya, serta mengevaluasi kinerja *Naïve Bayes* menggunakan metode pembagian data (*split ratio*) dan validasi silang (*cross validation*) untuk meningkatkan akurasi. Selain itu, perlu ditingkatkan banyaknya data yang diperoleh guna meningkatkan hasil akurasi.

REFERENCES

- [1] U. Kulsum, M. Jajuli, and N. Sulistiyowati, "Analisis Sentimen Aplikasi WETV di Google Play Store Menggunakan Algoritma Support Vector Machine," *J. Appl. Informatics Comput.*, vol. 6, no. 2, pp. 205–212, 2022, doi: 10.30871/jaic.v6i2.4802.
- [2] M. Fudhail Ferio Supeli and S. Setiaji, "Klasifikasi Sentimen Positif Dan Negatif Pada Aplikasi Vidio Dengan Algoritma K-Nearest Neighbor," *Indones. J. Comput. Sci.*, vol. 2, no. 1, pp. 7–15, Apr. 2023, doi: 10.31294/ijcs.v2i1.1874.
- [3] H. A. Ramadhan, B. M. Purwaamijaya, and G. Guntara, "JTIM : Jurnal Teknologi Informasi dan Multimedia Pengaruh User Experience terhadap Customer Satisfaction pada Aplikasi Seluler Streaming Vidio," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 5, no. 2, pp. 122–133, 2023, doi: <https://doi.org/10.35746/jtim.v5i2.367>.
- [4] S. S. M. D. Wijaya, N. L. R. Purnawan, and N. M. R. A. Gelgel, "Motif dan Kepuasan Generasi Z Kota Denpasar Menggunakan Vidio sebagai Aplikasi Video Streaming," *J. Mediu.*, vol. 3, no. 1, pp. 34–41, 2022.
- [5] H. K. Rifqi *et al.*, "ANALISIS DESAIN USER INTERFACE," *Askara J. Seni dan Desain*, vol. 2, no. 128, pp. 98–108, 2023, doi: <https://doi.org/10.20895/askara.v2i2.1129>.
- [6] F. N. H. Meliyawati, "Analisis Sentimen Pengguna Aplikasi CapCut Pada Ulasan di Play Store Menggunakan Metode Naïve Bayes," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 4, pp. 2272–2280, 2024, doi: 10.30865/klik.v4i4.1555.
- [7] P. Aditiya, U. Enri, and I. Maulana, "Analisis Sentimen Ulasan Pengguna Aplikasi Myim3 Pada Situs Google Play Menggunakan Support Vector Machine," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, p. 1020, Aug. 2022, doi: 10.30865/jurikom.v9i4.4673.
- [8] R. A. Saputra *et al.*, "Analisis Sentimen Aplikasi Tokocrypto Berdasarkan Ulasan Pada Google Play Store Menggunakan Metode Naïve Bayes," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 4, pp. 2028–2036, 2024, doi: 10.30865/klik.v4i4.1707.
- [9] D. R. Alghifari, M. Edi, and L. Firmansyah, "Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia," *J. Manaj. Inform.*, vol. 12, no. 2, pp. 89–99, 2022, doi: 10.34010/jamika.v12i2.7764.
- [10] M. Iqbal, A. D. Wiranata, R. Suwito, R. F. Ananda, U. Muhammadiyah, and P. Hamka, "Perbandingan Algoritma Naïve Bayes, KNN, dan Decision Tree terhadap Ulasan Aplikasi Threads dan Twitter," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 3, pp. 1799–1807, 2023, doi: 10.30865/klik.v4i3.1402.
- [11] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *J. KomtekInfo*, vol. 10, pp. 1–7, Jan. 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [12] N. Q. Rizkina and F. N. Hasan, "Analisis Sentimen Komentar Netizen Terhadap Pembubaran Konser NCT 127 Menggunakan Metode Naïve Bayes," *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1136–1144, Jul. 2023, doi: 10.47065/josh.v4i4.3803.
- [13] H. Z. Muflih, A. R. Abdillah, and F. N. Hasan, "Analisis Sentimen Ulasan Pengguna Aplikasi Ajaib Menggunakan Metode Naïve Bayes," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 3, pp. 1613–1621, 2023, doi: 10.30865/klik.v4i3.1303.
- [14] T. A. Sari, E. Sinduningrum, and F. Noor Hasan, "KLIK: Kajian Ilmiah Informatika dan Komputer Analisis Sentimen Ulasan Pelanggan Pada Aplikasi Fore Coffee Menggunakan Metode Naïve Bayes," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 3, no. 6, pp. 773–779, 2023, doi: 10.30865/klik.v3i6.884.
- [15] M. T. Nitamia and H. Februariyanti, "Analisis Sentimen Ulasan Ekpedisi J&T Expres Menggunakan Algoritma Naive Bayes," *J. Manaj. Inform. Sist. Inf.*, vol. 5, no. 1, pp. 20–29, 2022, doi: <https://doi.org/10.36595/misi.v5i1.396>.
- [16] S. A. R. Rizaldi, S. Alam, and I. Kurniawan, "Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes," *STORAGE J. Ilm. Tek. dan Ilmu Komput.*, vol. 2, no. 3, pp. 109–117, Aug. 2023, doi: 10.55123/storage.v2i3.2334.
- [17] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *SMATIKA J.*, vol. 10, no. 02, pp. 71–76, Dec. 2020, doi: 10.32664/smatika.v10i02.455.
- [18] N. Lestari, E. Haerani, and R. M. Candra, "Analisa Sentimen Ulasan Aplikasi Wetv Untuk Peningkatan Layanan Menggunakan Metode Naïve Bayes," *J. Inf. Syst. Res.*, vol. 4, no. 3, pp. 874–882, Apr. 2023, doi: 10.47065/josh.v4i3.3355.
- [19] W. Parasati, F. A. Bachtiar, and N. Y. Setiawan, "Analisis Sentimen Berbasis Aspek pada Ulasan Pelanggan Restoran Bakso President Malang dengan Metode Naïve Bayes Classifier," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 4, pp. 1090–1099, 2020.
- [20] H. P. Doloksaribu and Y. T. Samuel, "KOMPARASI ALGORITMA DATA MINING UNTUK ANALISIS SENTIMEN APLIKASI PEDULILINDUNGI," *J. Teknol. Inf. J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 16, no. 1, pp. 1–11, 2022, doi: <https://doi.org/10.47111/jti.v16i1.3747>.

- [21] F. P. Nursyamsyi and F. N. Hasan, "Klasifikasi Sentimen Terhadap Aplikasi Identitas Kependudukan Digital Menggunakan Algoritma Naïve Bayes dan SVM," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 3, pp. 1788–1798, 2023, doi: 10.30865/klik.v4i3.1517.
- [22] F. Septianingrum, J. H. Jaman, and U. Enri, "Analisis Sentimen Pada Isu Vaksin Covid-19 di Indonesia dengan Metode Naive Bayes Classifier," *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 4, p. 1431, Oct. 2021, doi: 10.30865/mib.v5i4.3260.
- [23] H. Dhery, A. Assyam, and F. N. Hasan, "Analisis Sentimen Twitter Terhadap Perpindahan Ibu Kota Negara Ke IKN Nusantara Menggunakan Orange Data Mining," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 1, pp. 341–349, 2023, doi: 10.30865/klik.v4i1.957.
- [24] O. I. Gifari, M. Adha, I. R. Hendrawan, F. Freddy, and S. Durrand, "Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine," *JIFOTECH (JOURNAL Inf. Technol.*, vol. 2, no. 1, pp. 36–40, 2022.
- [25] A. Saepulrohman, S. Saepudin, and D. Gustian, "Analisis Sentimen Kepuasan Pengguna Aplikasi Whatsapp Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine," *is Best Account. Inf. Syst. Inf. Technol. Bus. Enterp. this is link OJS usf@*, vol. 6, no. 2, pp. 91–105, Dec. 2021, doi: 10.34010/aisthebest.v6i2.4919.
- [26] I. P. Rahayu, A. Fauzi, and J. Indra, "Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Dan Support Vector Machine," *J. Sist. Komput. dan Inform.*, vol. 4, no. 2, p. 296, Dec. 2022, doi: 10.30865/json.v4i2.5381.
- [27] W. A. Prabowo and C. Wiguna, "Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM," *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 1, p. 149, Jan. 2021, doi: 10.30865/mib.v5i1.2604.
- [28] S. I. Nurhafida, F. Sembiring, J. Raya, and C. No, "Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)," *J. Sains Komput. Inform.*, vol. 6, pp. 317–327, 2022, doi: <http://dx.doi.org/10.30645/j-sakti.v6i1.447>.
- [29] 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.
- [30] N. P. Husain, S. Sukirman, and S. SAJIAH, "Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine," *J. Syst. Comput. Eng.*, vol. 5, no. 1, pp. 91–102, Jan. 2024, doi: 10.61628/jsce.v5i1.1105.
- [31] I. Verawati and B. S. Audit, "Algoritma Naïve Bayes Classifier Untuk Analisis Sentiment Pengguna Twitter Terhadap Provider By.u," *J. MEDIA Inform. BUDIDARMA*, vol. 6, no. 3, p. 1411, Jul. 2022, doi: 10.30865/mib.v6i3.4132.
- [32] K. Kusuma Dewi, I. Kaniawulan, and C. Dewi Lestari, "ANALISIS SENTIMEN PENGGUNA APLIKASI JAMSOSTEK MOBILE (JMO) PADA APPSTORE MENGGUNAKAN METODE NAIVE BAYES," *Simtek J. Sist. Inf. dan Tek. Komput.*, vol. 8, no. 2, pp. 333–338, Oct. 2023, doi: 10.51876/simtek.v8i2.286.
- [33] A. D. Adhi Putra, "Analisis Sentimen pada Ulasan pengguna Aplikasi Bibit Dan Bareksa dengan Algoritma KNN," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 2, pp. 636–646, Jun. 2021, doi: 10.35957/jatisi.v8i2.962.
- [34] Irvandi, B. Irawan, and O. Nurdiawan, "NAIVE BAYES DAN WORDCLOUD UNTUK ANALISIS SENTIMEN WISATA HALAL PULAU LOMBOK," *INFOTECH J.*, vol. 9, no. 1, pp. 236–242, May 2023, doi: 10.31949/infotech.v9i1.5322.