

ANALISIS SENTIMEN TERHADAP ULASAN APLIKASI MEDIA SOSIAL DI GOOGLE PLAY MENGGUNAKAN ALGORITMA NAIVE BAYES

Maksum Zainuddin Siregar, Andi Marwan Elhanafi, Dedy Irwan

Sistem Informasi, Universitas Harapan Medan

Jl. H.M. Jhoni No. 70 C Medan

Maxsiregar7z@gmail.com

ABSTRAK

Di era digital saat ini, media sosial seperti Instagram dan TikTok memiliki jutaan pengguna yang aktif memberikan ulasan di platform Google Play Store. Ulasan pengguna ini merupakan sumber data berharga yang dapat membantu pengembang memahami sentimen dan persepsi pengguna. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna terhadap aplikasi Instagram dan TikTok di Google Play Store menggunakan algoritma Naive Bayes. Metode ini dipilih karena kesederhanaan dan efisiensinya dalam klasifikasi teks. Penelitian ini menunjukkan bahwa sebagian besar ulasan pengguna dikategorikan sebagai sentimen negatif dengan persentase mencapai 87,5%, mengindikasikan ketidakpuasan yang dominan di antara pengguna. Namun, hasil analisis juga mengungkapkan keterbatasan algoritma Naive Bayes dalam menangani variasi bahasa informal dan konteks spesifik dari ulasan. Model cenderung memberikan bias ke arah prediksi negatif meskipun terdapat ulasan yang seharusnya diklasifikasikan sebagai positif atau netral.

Kata Kunci : Analisis Sentimen, Naive Bayes, Google Play, Sosial Media

1. PENDAHULUAN

Di era digital saat ini, aplikasi media sosial seperti Instagram dan TikTok telah menjadi bagian penting dari kehidupan sehari-hari, dengan jutaan pengguna yang aktif berinteraksi dan berbagi konten. Kedua platform ini tidak hanya populer di kalangan pengguna, tetapi juga memiliki aplikasi yang tersedia di Google Play Store di mana pengguna dapat memberikan ulasan dan feedback terkait pengalaman mereka. Ulasan ini menjadi sumber data berharga bagi pengembang dan pemasar untuk memahami persepsi pengguna terhadap aplikasi mereka. Google Play Store adalah platform utama untuk distribusi aplikasi di perangkat Android, dan ulasan yang diberikan oleh pengguna di sini mencerminkan pengalaman mereka dalam menggunakan aplikasi tersebut. Ulasan dari aplikasi Instagram dan TikTok di Google Play Store menawarkan wawasan yang signifikan mengenai bagaimana pengguna menilai aplikasi media sosial ini, baik dari segi fitur, kinerja, maupun kepuasan keseluruhan. Analisis sentimen dari ulasan ini dapat memberikan gambaran yang jelas tentang pandangan pengguna dan area yang mungkin memerlukan perbaikan [1].

Namun, menganalisis sentimen dari ulasan yang sangat banyak dan beragam memerlukan metode yang efisien dan otomatis. Salah satu pendekatan yang dapat digunakan adalah algoritma Naive Bayes, yang merupakan teknik klasifikasi berbasis probabilitas. Algoritma ini dapat membantu dalam mengklasifikasikan ulasan menjadi kategori sentimen yang berbeda, seperti positif, negatif, atau netral, berdasarkan frekuensi kata-kata yang muncul dalam ulasan tersebut. Meskipun Naive Bayes dikenal karena kesederhanaannya dan efisiensinya, analisis sentimen dari data ulasan memerlukan penanganan khusus terhadap variasi bahasa dan konteks. Ulasan pada

Google Play Store sering kali mengandung bahasa informal dan istilah yang spesifik, yang dapat mempengaruhi akurasi model. Oleh karena itu, preprocessing data dan penyesuaian algoritma mungkin diperlukan untuk meningkatkan hasil analisis.

Media sosial adalah platform digital yang memungkinkan pengguna untuk membuat, berbagi, dan berinteraksi dengan konten serta membangun hubungan dengan pengguna lain di seluruh dunia. Pada dasarnya, media sosial memberikan ruang bagi individu untuk mengekspresikan diri, berkomunikasi, dan terlibat dalam komunitas online, yang mencakup berbagai bentuk konten seperti teks, gambar, video, dan lain-lain. Dengan perkembangan teknologi dan peningkatan akses internet, media sosial telah menjadi bagian integral dari kehidupan sehari-hari, mempengaruhi cara orang berkomunikasi, memperoleh informasi, dan berinteraksi satu sama lain [2].

Sementara Instagram dan TikTok menawarkan pengalaman pengguna yang berbeda, baik dalam bentuk konten maupun interaksi, analisis sentimen dapat memberikan wawasan yang komprehensif tentang bagaimana pengguna merasakan aplikasi mereka. Instagram berfokus pada konten visual dengan caption dan komentar, sedangkan TikTok menekankan video pendek dengan teks overlay. Memahami sentimen di kedua platform ini dapat membantu pengembang untuk menyesuaikan fitur dan strategi pemasaran mereka lebih efektif. Dalam konteks analisis ini, algoritma Naive Bayes akan diterapkan pada ulasan aplikasi Instagram dan TikTok di Google Play Store. Dengan memanfaatkan algoritma ini, diharapkan dapat mengidentifikasi pola sentimen yang mendasari ulasan pengguna dan memberikan wawasan yang berharga untuk pengembangan dan perbaikan aplikasi.

Kesimpulannya, analisis sentimen terhadap ulasan aplikasi media sosial di Google Play menggunakan algoritma Naive Bayes menyediakan metode yang praktis untuk memahami pengalaman pengguna. Dengan fokus pada data ulasan dari Instagram dan TikTok, penelitian ini bertujuan untuk memberikan pemahaman yang lebih baik tentang sentimen pengguna dan membantu dalam meningkatkan kualitas aplikasi serta strategi pemasaran di pasar yang sangat kompetitif [3].

Adapun penelitian terdahulu yang dilakukan oleh [4] penelitian ini menghasilkan sebuah sistem analisis sentimen pada jasa ojek online menggunakan metode Naive Bayes. Proses penelitian ini menggunakan masukkan berupa tweet mention jasa ojek online di Twitter. Hasil dari pengujian, sistem mampu mengklasifikasi sentimen menggunakan Naive Bayes dengan akurasi yang dihasilkan sebesar 80% berdasarkan 800 data tweet yang terdiri atas 300 data latih dan 500 data uji. Terdapat kesalahan pada data uji fitur yang muncul tidak sesuai dengan klasifikasinya. Akurasi klasifikasi dapat ditingkatkan dengan menambah jumlah data latih.

2. TINJAUAN PUSTAKA

Tinjauan pustaka berisi kajian teoritis, penelitian terdahulu, dan landasan konseptual yang relevan dengan topik penelitian. Bagian ini bertujuan untuk memberikan gambaran umum mengenai penelitian terkait, menjelaskan konsep atau teori yang digunakan

2.1. Analisis Sentimen

Analisis sentimen adalah proses menilai sikap, pendapat, atau emosi yang diekspresikan dalam teks. Biasanya, ini dilakukan untuk menentukan apakah sentimen yang terkandung dalam suatu teks positif, negatif, atau netral. Dalam konteks digital, analisis sentimen banyak digunakan untuk memahami bagaimana perasaan publik terhadap produk, layanan, atau berbagai topik yang sedang tren. Proses ini memanfaatkan teknik pemrosesan bahasa alami (NLP) dan pembelajaran mesin untuk menganalisis data tekstual yang besar dan heterogen, seperti ulasan pelanggan, posting media sosial, atau artikel berita. Dalam era digital saat ini, di mana volume data teks yang dihasilkan oleh pengguna terus meningkat, analisis sentimen menjadi alat penting bagi bisnis dan peneliti untuk memahami persepsi masyarakat secara luas. Misalnya, ulasan pelanggan di platform e-commerce, komentar di media sosial, atau feedback dari survei online dapat dianalisis untuk mendapatkan wawasan mengenai kepuasan pelanggan atau kinerja produk. Wawasan ini dapat membantu perusahaan untuk mengambil keputusan strategis, seperti perbaikan produk atau penyesuaian strategi pemasaran [5].

2.2. Data Ulasan

Data ulasan merujuk pada informasi yang dihasilkan oleh pengguna ketika mereka memberikan

pendapat atau evaluasi terhadap suatu produk, layanan, atau pengalaman. Ulasan ini biasanya berbentuk teks yang disertai dengan penilaian numerik, seperti bintang atau skor, yang mencerminkan tingkat kepuasan pengguna. Dalam era digital, data ulasan sering ditemukan di berbagai platform online, termasuk situs e-commerce, aplikasi, media sosial, dan forum diskusi. Data ulasan memainkan peran penting dalam membantu calon konsumen membuat keputusan pembelian yang lebih informasi dan membantu bisnis memahami persepsi pengguna terhadap produk atau layanan mereka [6].

2.3. Google Play

Google Play, sebelumnya dikenal sebagai Android Market, adalah toko aplikasi resmi yang dikembangkan dan dikelola oleh Google untuk perangkat yang menjalankan sistem operasi Android. Diluncurkan pada tahun 2008, Google Play tidak hanya menyediakan aplikasi, tetapi juga menawarkan berbagai konten digital lainnya seperti musik, buku, film, dan perangkat keras tertentu. Google Play telah berkembang menjadi platform distribusi konten digital terbesar di dunia, yang digunakan oleh jutaan pengembang dan miliaran pengguna di seluruh dunia. Dengan lebih dari tiga juta aplikasi dan permainan yang tersedia, Google Play menjadi pusat ekosistem Android yang sangat vital [7].

2.4. Aplikasi Tiktok

TikTok adalah aplikasi media sosial yang memungkinkan pengguna untuk membuat, berbagi, dan menonton video pendek yang biasanya diiringi dengan musik. Diluncurkan pada tahun 2016 oleh perusahaan teknologi asal China, ByteDance, TikTok awalnya dikenal dengan nama Douyin di China. Versi internasionalnya kemudian diluncurkan dengan nama TikTok pada tahun 2017. Dalam waktu singkat, TikTok berhasil menarik perhatian global dan menjadi salah satu platform media sosial dengan pertumbuhan tercepat di dunia. Aplikasi ini menawarkan format video yang unik dan interaktif, yang sangat disukai oleh pengguna muda di seluruh dunia. TikTok menonjol di antara aplikasi media sosial lainnya karena format videonya yang pendek, berdurasi 15 hingga 60 detik, yang memungkinkan pengguna untuk dengan cepat membuat konten yang kreatif dan menarik. Pengguna dapat menambahkan berbagai efek visual, filter, dan suara latar untuk memperkaya video mereka. Fitur-fitur ini memungkinkan ekspresi diri yang lebih bebas dan sering kali menghasilkan tren dan tantangan viral yang cepat menyebar di kalangan pengguna. Tren ini sering kali didorong oleh tagar (#) dan musik tertentu yang menjadi populer di kalangan komunitas TikTok [8].

2.5. Aplikasi Instagram

Instagram adalah sebuah aplikasi media sosial yang memungkinkan pengguna untuk berbagi foto dan video, serta berinteraksi dengan konten yang diposting

oleh orang lain. Aplikasi ini diluncurkan pada Oktober 2010 oleh Kevin Systrom dan Mike Krieger. Awalnya, Instagram dirancang sebagai aplikasi berbagi foto dengan fitur filter yang memungkinkan pengguna untuk mengedit gambar mereka sebelum membagikannya. Keunikan fitur filter ini, yang memberikan sentuhan estetik pada foto dengan cepat, menjadi salah satu faktor yang membuat Instagram sangat populer dalam waktu singkat. Pada tahun 2012, Instagram diakuisisi oleh Facebook dengan nilai sekitar \$1 miliar, sebuah langkah yang menandai pentingnya Instagram dalam lanskap media sosial global. Setelah akuisisi, Instagram mengalami pertumbuhan yang luar biasa, baik dari segi jumlah pengguna maupun fitur-fitur yang ditawarkannya. Fitur-fitur baru seperti Instagram Stories, IGTV, dan Instagram Shopping terus diperkenalkan, yang memungkinkan pengguna untuk berbagi konten dengan cara yang lebih dinamis dan interaktif. Stories, yang memungkinkan pengguna untuk membagikan foto dan video yang hilang setelah 24 jam, terinspirasi oleh Snapchat dan menjadi salah satu fitur paling populer di Instagram [4].

2.6. Naive Bayes

Naive Bayes merupakan salah satu algoritma dari machine learning. Dalam perkembangan database, Naive Bayes termasuk supervised learning yaitu jenis machine learning yang membutuhkan sampel sebagai data latih yang memiliki label. Supervised learning dikelompokkan menjadi dua yaitu klasifikasi dan regresi. Klasifikasi pada saat variabel menjadi kategori seperti merah atau kuning, penyakit atau tidak ada penyakit, dan sebagainya. Regresi pada saat variabel berupa nilai riil seperti berat, nilai uang, dan sebagainya. Naive Bayes termasuk supervised learning klasifikasi seperti contoh lainnya yaitu Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Artificial Neural Network (ANN), Trees Gradient Boosted (TGB), dan Random Trees (RT) sedangkan regresi seperti Decision Tree, Logistic Regression, dan Kernel Regression (Roihan et al., 2020). Metode ini merupakan bagian dari metode bayes yang digunakan dalam klasifikasi teks didasarkan pada model penyederhanaan bahwa nilai atribut secara kondisional bersifat bebas jika nilai output diberikan [4].

$$P = (Y|X) = \frac{(x|y)(Y)}{P(X)} \quad (1)$$

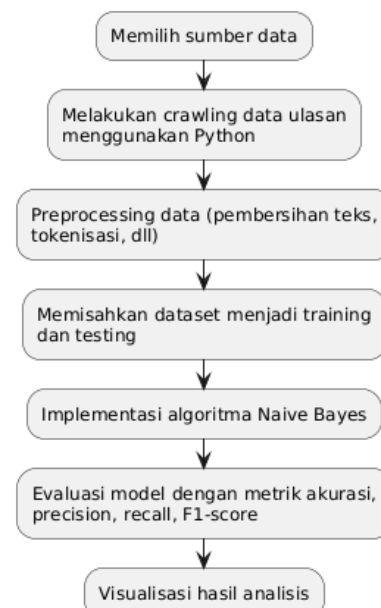
Persamaan di atas menunjukkan bahwa Y sebagai kelas spesifik, sedangkan X sebagai kelas yang belum diketahui. $(Y|X)$ merupakan probabilitas dari kelas berdasarkan hipotesa sebelumnya, sedangkan $P(X)$ merupakan probabilitas dari Y. $20 P(Y|X)$ adalah hasil perkalian antara likelihood dan prior dibagi evidence. Likelihood adalah probabilitas atribut data X pada kelas Y, prior adalah probabilitas kelas Y dari total data set, dan evidence adalah probabilitas atribut data X dari total data set [9].

3. METODE PENELITIAN

Metode penelitian menjelaskan pendekatan, desain, teknik pengumpulan data, dan analisis yang digunakan dalam penelitian. Bagian ini bertujuan untuk memastikan bahwa penelitian dilakukan secara sistematis dan dapat dipertanggungjawabkan.

3.1. Alur Penelitian

Bagian ini menjelaskan secara rinci tahapan-tahapan yang dilakukan dalam penelitian ini, mulai dari pemilihan dan pengumpulan data, hingga analisis dan evaluasi hasil menggunakan algoritma Naive Bayes. Setiap langkah disusun secara sistematis untuk memastikan akurasi dan relevansi hasil yang diperoleh.



Gambar 1. Alur Penelitian

Adapun keterangan gambar 1. yaitu Penelitian ini bertujuan untuk menganalisis sentimen terhadap ulasan aplikasi media sosial yang tersedia di Google Play dengan menggunakan algoritma Naive Bayes. Proses penelitian dimulai dengan pemilihan sumber data, di mana ulasan yang akan dianalisis berasal dari aplikasi media sosial seperti Instagram dan TikTok, yang ditulis dalam Bahasa Indonesia. Langkah berikutnya adalah melakukan crawling data ulasan tersebut menggunakan skrip Python. Dalam proses ini, data seperti teks ulasan, rating, dan informasi relevan lainnya dikumpulkan dari Google Play. Setelah data berhasil dikumpulkan, dilakukan tahap preprocessing yang meliputi pembersihan teks, tokenisasi, dan normalisasi data untuk memastikan kualitas data yang akan digunakan dalam analisis. Data yang telah diproses kemudian dibagi menjadi dua set: dataset untuk training dan dataset untuk testing. Pembagian ini penting untuk membangun dan menguji model klasifikasi yang akan digunakan. Algoritma Naive Bayes diterapkan pada data training untuk membangun model analisis sentimen.

Setelah model dibangun, dilakukan evaluasi menggunakan beberapa metrik seperti akurasi, precision, recall, dan F1-score. Evaluasi ini bertujuan untuk mengukur kinerja model dalam mengklasifikasikan sentimen ulasan sebagai positif, negatif, atau netral. Terakhir, hasil analisis dan evaluasi disajikan dalam bentuk visualisasi yang dapat berupa grafik atau tabel untuk mempermudah interpretasi hasil penelitian.

4. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penelitian yang diperoleh dan membahasnya secara mendalam.

4.1. Crawling Data

Gambar di bawah ini menunjukkan hasil dari proses crawling data yang dilakukan untuk mengumpulkan ulasan dari aplikasi media sosial di Google Play. Proses crawling ini menggunakan skrip Python untuk secara otomatis mengekstrak data ulasan pengguna yang mencakup konten ulasan, skor yang diberikan pengguna, dan jumlah "thumbs up" atau likes yang diterima oleh ulasan tersebut.

	content	score	thumbsUpCount
0	semakin update semakin banyak bug, saat menont...	3	36
1	Sekarang bugnya jadi makin banyak dan cukup me...	2	844
2	Saya suka instagram begitu banyak video bermanf...	4	126
3	Jelek banget, bug dari tahun kemarin belum di ...	1	505
4	terjadi lagi! setiap upload story selalu gabis...	1	54

Gambar 2. Proses Crawling Data Instagram

Pada gambar 2 Proses crawling data adalah langkah awal yang krusial dalam penelitian ini, di mana data ulasan pengguna dari aplikasi media sosial diambil langsung dari Google Play. Data yang berhasil dikumpulkan ditampilkan dalam bentuk tabel yang berisi beberapa kolom penting. Kolom "content" berisi teks ulasan yang diberikan oleh pengguna. Kolom "score" menunjukkan rating yang diberikan oleh pengguna terhadap aplikasi, dengan skala 1 hingga 5, di mana 1 menunjukkan ketidakpuasan dan 5 menunjukkan kepuasan maksimal. Kolom "thumbsUpCount" menunjukkan berapa banyak pengguna lain yang menyukai atau setuju dengan ulasan tersebut.

	content	score	thumbsUpCount
0	Akun saya terblokir sudah dua kali . akun pert...	1	177
1	kenapa foto yang di simpan malah jadi video dn...	1	512
2	Pihak tiktok kami sangat berterima kasih atas ...	3	107
3	Wah! keren apk ini sangat bagus utk org yg bet...	5	875
4	Knp pas lagi nonton tiktok tiba tiba tiba ada ...	2	11

reviewCreatedVersion	at replyContent	repliedAt	appVersion		
0	35.3.4	2024-08-04 01:02:25	None	NaT	35.3.4
1	35.6.3	2024-07-20 09:06:35	None	NaT	35.6.3
2	35.3.4	2024-07-08 23:01:30	None	NaT	35.3.4
3	35.3.4	2024-07-02 08:57:09	None	NaT	35.3.4
4	35.8.3	2024-08-01 05:28:09	None	NaT	35.8.3

Number of reviews scraped: 508

Gambar 3. Crawling Data TikTok

4.2. Explore

Hasil crawling terdiri dari beberapa kolom yang berisi informasi penting dari ulasan aplikasi di Google Play. Kolom content berisi teks ulasan yang ditulis oleh pengguna. Score merupakan skor atau rating yang diberikan oleh pengguna, dengan skala antara 1 hingga 5, di mana 1 menunjukkan ulasan yang sangat negatif

dan 5 menunjukkan ulasan yang sangat positif. Kolom thumbsUpCount mencerminkan jumlah "thumbs up" atau likes yang diterima oleh ulasan tersebut dari pengguna lain. Untuk keperluan analisis sentimen, atribut yang paling relevan dan penting adalah content yang berisi teks ulasan. Oleh karena itu, fokus analisis akan dilakukan pada teks ulasan tersebut untuk mengidentifikasi sentimen pengguna terhadap aplikasi. Kolom lain seperti score dan thumbsUpCount dapat memberikan konteks tambahan dalam analisis, namun penekanan utama adalah pada konten teks yang akan diolah lebih lanjut.

userName	score	at	content
Gilbert Paef	1	2024-08-11 07:48:01	apaan ni mo pos foto aja gak bisa jangan kasi balasan dong mau pos gak bisa... aku gak mau tahu harus perbaiki akun gue gak bisa lagi buat apot foto dan video
Leey Chi	3	2024-08-10 16:41:14	Tolong di perbaiki, kenapa sering kuar sendiri
Clarita Da Costa	1	2024-08-09 22:32:30	Kenapa aplikasi lain ada fitur Sulap kok saya punya ini nggak ada sih... padahal dulu itu ada terus hilang kenapa yeh?
Fibri Sayang	1	2024-08-08 23:26:54	Kenpa tiktok saya di blokir live permanen, sedangkan saya ga melanggar apapun, sudah 4 tahun tiktok saya ga bisa live
Adryel Dosantos	5	2024-08-08 13:54:09	Aplikasi sangat membantu belajar Menghibur dan menghilangkan bosang
Deva Sri mardiana	5	2024-08-08 06:06:58	Kenapa tiktok kami satu cepet rusak
Eva Lopes	5	2024-08-07 11:44:03	Sangat berguna
Fatma Masqula	1	2024-08-07 11:34:13	Ini itu selalu keluar iklan paze!

Gambar 4. Explore

Penjelasan gambar 4. yaitu Gambar di atas menunjukkan hasil dari proses crawling data yang dilakukan untuk mengumpulkan ulasan pengguna terhadap aplikasi media sosial di Google Play. Setiap entri dalam tabel mencakup beberapa informasi penting yang diperoleh dari ulasan tersebut, termasuk nama pengguna, skor yang diberikan, tanggal dan waktu ulasan dibuat, serta isi dari ulasan itu sendiri.

4.3. Modify

Modifikasi dilakukan pada dataset ulasan aplikasi media sosial di Google Play melalui tahapan text preprocessing. Tahapan ini bertujuan agar sistem dapat lebih mengenali dan memahami bentuk data yang akan dianalisis. Dengan melakukan preprocessing, data yang tidak terstruktur dapat diubah menjadi lebih terorganisir dan siap untuk dianalisis. Dalam text preprocessing terdapat beberapa proses penting yang dilakukan, yaitu cleansing (pembersihan teks dari karakter yang tidak diperlukan), tokenizing (memecah teks menjadi kata-kata terpisah), case folding (mengubah semua teks menjadi huruf kecil untuk konsistensi), dan stopword removal (menghapus kata-kata umum yang tidak memiliki nilai dalam analisis). Setiap proses ini dirancang untuk memastikan bahwa data yang digunakan dalam analisis sentimen berkualitas tinggi dan dapat memberikan hasil yang akurat.

4.4. Cleaning

Pada proses ini, dilakukan pembersihan teks dengan menghapus elemen-elemen yang tidak dibutuhkan dan tidak memiliki makna signifikan untuk analisis sentimen, seperti emoji, tanda baca berlebihan, angka, dan spasi berlebih. Selain itu, dilakukan juga penghapusan kata-kata umum yang tidak relevan dalam konteks analisis, seperti kata penghubung dan

preposisi. Berikut ini adalah contoh hasil cleansing pada salah satu dataset ulasan.

```
Setelah Cleansing:
content \
0 apaan ni mo pos foto aja gak bisa jangan kasi ...
1 Tolong di perbaiki kenapa sering kluar sendiri
2 Kenapa aplikasi lain ada fitur Sulap kok saya ...
3 Kenapa tiktok saya di blokir live permanen, seda...
4 Aplikasi sangat membantu belajar Meng hiburan dan...

cleaned_content
0 apaan ni mo pos foto aja gak bisa jangan kasi ...
1 Tolong di perbaiki kenapa sering kluar sendiri
2 Kenapa aplikasi lain ada fitur Sulap kok saya ...
3 Kenapa tiktok saya di blokir live permanen seda...
4 Aplikasi sangat membantu belajar Meng hiburan dan...
```

Gambar 5. Setelah Proses Cleaning Data

Gambar 5 menampilkan hasil dari proses cleansing yang dilakukan pada data ulasan aplikasi media sosial. Proses cleansing adalah salah satu tahap penting dalam text preprocessing, yang bertujuan untuk membersihkan teks dari elemen-elemen yang tidak relevan atau mengganggu analisis lebih lanjut. Pada bagian atas gambar, terlihat data awal yang belum diproses di kolom `content`. Ulasan-ulasan ini mengandung berbagai karakter yang mungkin tidak diperlukan, seperti tanda baca, angka, atau spasi yang berlebihan. Misalnya, pada ulasan pertama terlihat penggunaan tanda baca yang berlebihan, dan ulasan lainnya mungkin memiliki kesalahan penulisan atau karakter yang tidak relevan. Setelah dilakukan proses cleansing, hasilnya ditampilkan pada kolom `cleaned_content` di bagian bawah gambar. Perubahan yang dilakukan meliputi:

- Penghapusan Karakter yang Tidak Diperlukan: Tanda baca yang berlebihan, angka, atau karakter lain yang tidak memberikan kontribusi pada analisis sentimen telah dihapus.
- Normalisasi Teks: Spasi yang berlebihan dihilangkan, sehingga teks menjadi lebih konsisten dan rapi.
- Penghapusan Tanda Baca Berlebihan: Misalnya, beberapa tanda baca seperti tanda seru atau tanda tanya yang berlebihan dihilangkan untuk memastikan bahwa teks lebih bersih dan siap untuk analisis.

4.5. Tokenisasi

Selanjutnya, proses tokenisasi dilakukan untuk memecah teks ulasan yang telah dibersihkan menjadi potongan kata-kata atau token. Proses ini bertujuan untuk memisahkan setiap kata dalam teks sehingga dapat dianalisis secara individual. Tokenisasi memungkinkan sistem untuk memahami komponen-komponen dari setiap kalimat dalam ulasan, yang kemudian dapat digunakan untuk analisis lebih lanjut, seperti penghitungan frekuensi kata atau penerapan model klasifikasi. Berikut ini adalah contoh hasil tokenisasi pada dataset ulasan.

```
Setelah Tokenizing:
cleaned_content \
0 apaan ni mo pos foto aja gak bisa jangan kasi ...
1 Tolong di perbaiki kenapa sering kluar sendiri
2 Kenapa aplikasi lain ada fitur Sulap kok saya ...
3 Kenapa tiktok saya di blokir live permanen seda...
4 Aplikasi sangat membantu belajar Meng hiburan dan...

tokenized_content
0 [apaan, ni, mo, pos, foto, aja, gak, bisa, jan...
1 [Tolong, di, perbaiki, kenapa, sering, kluar, ...
2 [Kenapa, aplikasi, lain, ada, fitur, Sulap, ko...
3 [Kenapa, tiktok, saya, di, blokir, live, perman...
4 [Aplikasi, sangat, membantu, belajar, Meng, hib...
```

Gambar 6. Hasil Tokenized

Gambar 6 menunjukkan hasil dari proses tokenisasi yang dilakukan pada teks ulasan yang telah melalui tahap cleansing. Proses tokenisasi merupakan langkah penting dalam text preprocessing, di mana teks yang telah dibersihkan dipecah menjadi kata-kata individu atau token. Setiap token ini dapat dianalisis secara terpisah dalam tahapan selanjutnya, seperti dalam penghitungan frekuensi kata atau untuk input ke dalam model pembelajaran mesin. Pada bagian atas gambar, terlihat kolom `cleaned_content` yang berisi teks ulasan setelah melalui proses pembersihan. Teks ini masih berupa kalimat yang utuh. Misalnya, salah satu contoh ulasan berbunyi: "Kenapa tiktok saya di blokir live permanen, sedangkan saya ga melanggar apapun". Setelah diterapkan proses tokenisasi, hasilnya ditampilkan di kolom `tokenized_content` pada bagian bawah gambar.

4.6. Case Folding

Proses case folding adalah langkah dalam text preprocessing yang bertujuan untuk mengubah semua huruf dalam teks ulasan menjadi huruf kecil (lowercase). Langkah ini penting untuk memastikan konsistensi dalam analisis teks, terutama dalam menghindari perbedaan antara kata yang sebetulnya identik namun ditulis dengan huruf kapital yang berbeda, seperti "Aplikasi" dan "aplikasi". Dengan melakukan case folding, kita dapat menyederhanakan teks dan mengurangi variasi yang tidak perlu, sehingga analisis sentimen dapat dilakukan dengan lebih akurat. Berikut adalah contoh hasil case folding pada dataset ulasan.

```
Setelah Case Folding:
cleaned_content \
0 apaan ni mo pos foto aja gak bisa jangan kasi ...
1 Tolong di perbaiki kenapa sering kluar sendiri
2 Kenapa aplikasi lain ada fitur Sulap kok saya ...
3 Kenapa tiktok saya di blokir live permanen seda...
4 Aplikasi sangat membantu belajar Meng hiburan dan...

lowercase_content
0 apaan ni mo pos foto aja gak bisa jangan kasi ...
1 tolong di perbaiki kenapa sering kluar sendiri
2 kenapa aplikasi lain ada fitur sulap kok saya ...
3 kenapa tiktok saya di blokir live permanen seda...
4 aplikasi sangat membantu belajar meng hiburan dan...
```

Gambar 7. Hasil Case Folding

Gambar 7 menunjukkan hasil dari proses case folding yang dilakukan pada teks ulasan setelah melalui tahap cleansing. Case folding adalah proses mengubah seluruh huruf dalam teks menjadi huruf

kecil (lowercase) untuk memastikan konsistensi dalam data. Proses ini sangat penting dalam text preprocessing, terutama ketika kita ingin memastikan bahwa kata-kata yang seharusnya sama tidak dianggap berbeda hanya karena perbedaan penggunaan huruf kapital. Pada bagian atas gambar, kolom `cleaned_content` menampilkan teks ulasan yang telah dibersihkan tetapi masih mempertahankan huruf besar dan kecil seperti pada aslinya. Misalnya, pada ulasan kelima terdapat kata "Aplikasi" dan "Menghibur" yang diawali dengan huruf kapital. Setelah dilakukan case folding, hasilnya ditampilkan di kolom `lowercase_content` pada bagian bawah gambar. Semua huruf kapital telah diubah menjadi huruf kecil, sehingga kata-kata seperti "Aplikasi" menjadi "aplikasi" dan "Menghibur" menjadi "menghibur".

4.7. Stopword Removal

Proses Stopword Removal adalah langkah penting dalam text preprocessing yang bertujuan untuk menghapus kata-kata umum yang sering muncul dalam teks tetapi memiliki sedikit atau tidak ada nilai informasi dalam analisis. Kata-kata seperti "dan", "yang", "untuk", dan sejenisnya dianggap sebagai stopwords, karena meskipun sering muncul, mereka tidak memberikan kontribusi yang signifikan dalam memahami makna atau sentimen dari teks. Dengan menghapus stopwords, kita dapat lebih fokus pada kata-kata yang benar-benar penting dan relevan dalam analisis sentimen. Berikut adalah contoh hasil Stopword Removal pada dataset ulasan.

```
Setelah Stopword Removal:

                                lowercase_content \
0  apaan ni mo pos foto aja gak bisa jangan kasi ...
1      tolong di perbaiki kenapa sering keluar sendiri
2  kenapa aplikasi lain ada fitur sulap kok saya ...
3  kenapa tiktok saya di blokir live permanen seda...
4  aplikasi sangat membantu belajar meng hiburan dan...

                                final_content
0  ni mo pos foto aja gak kasi batasan orang pos ...
1      tolong perbaiki kenapa sering keluar sendiri
2  aplikasi fitur sulap nggak sih teruss hilang yah
3  kenapa tiktok blokir live permanen ga melanggar...
4  aplikasi membantu belajar meng hiburan dang menhi...
```

Gambar 8. Hasil Stopword Removal

Gambar 8 menunjukkan hasil dari proses Stopword Removal yang dilakukan pada teks ulasan aplikasi setelah melalui tahapan cleansing, tokenizing, dan case folding. Stopword Removal adalah langkah penting dalam text preprocessing yang bertujuan untuk menghapus kata-kata yang sering muncul tetapi tidak memiliki kontribusi signifikan terhadap makna teks atau analisis sentimen, seperti "dan", "yang", "di", dan "untuk".

Berikut merupakan hasil keseluruhan preprosesing data yang telah dilakukan yaitu sebagai berikut :

						1 to 10 of 1000 entries	
username	score	at	content	cleaned_content	tokenized_content	lowercase_content	final_content
Gilbert Pael	1	2024-06-11 07:48:01	apain ni mo pos foto aja gak bisa jangan kasi bataasan orang mau pos gak bisa                                      				

Gambar 9. Hasil Preprocessing Data

Gambar 9 memberikan ilustrasi lengkap dari tahapan text preprocessing yang diterapkan pada dataset ulasan aplikasi media sosial, mencakup proses cleansing, tokenizing, case folding, dan stopword removal. Setiap kolom dalam tabel ini menunjukkan transformasi teks ulasan melalui setiap tahapan tersebut. Dimulai dengan kolom `userName`, kita melihat identitas pengguna yang memberikan ulasan, seperti "Gilbert Paef", yang memberikan ulasan dengan skor 1, menandakan ketidakpuasan yang tinggi. Waktu ulasan tersebut diberikan tercatat dalam kolom `at`, yang penting untuk analisis temporal.

4.8. Wordcloud

Wordcloud merupakan visualisasi data yang diambil dari kata-kata yang paling sering muncul dalam ulasan aplikasi di Google Play. Visualisasi ini memberikan gambaran visual tentang kata-kata yang dominan dalam teks ulasan, dengan ukuran kata yang lebih besar menunjukkan frekuensi kemunculan yang lebih tinggi.



Gambar 10. Wordcloud

Gambar 10 menampilkan hasil visualisasi wordcloud yang dihasilkan dari data ulasan aplikasi media sosial yang telah melalui proses text preprocessing. Wordcloud adalah representasi visual dari kata-kata yang paling sering muncul dalam teks ulasan, di mana ukuran setiap kata mencerminkan frekuensi kemunculannya. Proses dimulai dengan pengumpulan ulasan dari aplikasi TikTok dan Instagram, yang kemudian digabungkan menjadi satu dataset.

4.9. Data Latih dan Data Uji

Selanjutnya membagi data latih dan data uji yang akan dilakukan dalam penelitian ini yaitu sebagai berikut :

Jumlah data latih: 800
Jumlah data uji: 200

Contoh Data Latih:

	Content	Label
29	Mahatma Gandini	2
535	Xavier Fernandes	5
695	Alfhy Ady	2
557	rafily aminudin	2
836	Tanio Repal	5

Contoh Data Uji:

	Content	Label
521	alfarro reyyza	1
757	Martinah martinah	5
740	Tania Martins	1
660	Lizety Vidigal	1
411	RAMA DANUR AL-HAQ	5

Gambar 11. Data Uji dan Data Latih

Gambar 11 memberikan ilustrasi proses pembagian data ulasan yang telah diproses menjadi dua bagian utama: data latih (training data) dan data uji (testing data). Dari keseluruhan dataset, 80% dialokasikan sebagai data latih, yang terdiri dari 800 entri, sementara 20% sisanya dialokasikan sebagai data uji, yang terdiri dari 200 entri. Pembagian ini adalah langkah penting dalam persiapan data untuk melatih model pembelajaran mesin, seperti model Naive Bayes, yang sering digunakan untuk tugas analisis sentimen atau klasifikasi teks. Pada data latih, kolom "Content" berisi teks ulasan yang akan digunakan untuk melatih model.

4.10. Penerapan Metode Naive Bayes

Selanjutnya melakukan penerapan metode naive bayes menggunakan data uji dan data latih ada pun hasilnya sebagai berikut :

Classification Report:				
	precision	recall	f1-score	support
1	0.40	0.87	0.55	79
2	0.00	0.00	0.00	29
3	1.00	0.05	0.10	20
4	0.00	0.00	0.00	14
5	0.32	0.14	0.19	58
accuracy			0.39	200
macro avg	0.34	0.21	0.17	200
weighted avg	0.35	0.39	0.28	200

Gambar 12. Hasil Penerapan Metode Naive Bayes

Laporan klasifikasi yang ditampilkan dalam gambar memberikan gambaran tentang kinerja model Naive Bayes dalam mengklasifikasikan ulasan. Laporan ini mencakup beberapa metrik penting, seperti precision, recall, f1-score, dan support untuk setiap kelas, yang dalam hal ini adalah label 1 hingga 5. Precision menunjukkan seberapa tepat prediksi model untuk setiap kelas. Dari laporan ini, kita melihat bahwa precision untuk kelas 1 adalah 0.40, yang berarti 40% dari prediksi untuk kelas 1 benar. Namun, precision untuk kelas 2 dan 4 adalah 0.00, menunjukkan bahwa model gagal membuat prediksi yang benar untuk kelas-kelas ini. Precision untuk kelas 3 adalah 1.00, tetapi ini mungkin menunjukkan adanya overfitting atau jumlah prediksi yang sangat sedikit untuk kelas ini, sementara kelas 5 memiliki precision sebesar 0.32.

Recall mengukur seberapa baik model mampu menangkap semua contoh aktual dari sebuah kelas. Recall untuk kelas 1 adalah 0.87, menunjukkan bahwa model berhasil mengenali 87% dari semua data yang

seharusnya diklasifikasikan sebagai kelas 1. Namun, recall untuk kelas 2, 3, 4, dan 5 sangat rendah, dengan beberapa kelas memiliki nilai 0.00, yang berarti model tidak mampu menangkap contoh dari kelas-kelas tersebut dengan baik. F1-score adalah rata-rata harmonis dari precision dan recall, memberikan keseimbangan antara keduanya. F1-score untuk kelas 1 adalah 0.55, menunjukkan keseimbangan yang cukup baik antara precision dan recall untuk kelas ini. Namun, F1-score untuk kelas lainnya sangat rendah, menandakan kesulitan model dalam menyeimbangkan precision dan recall untuk kelas-kelas tersebut.

Support menunjukkan jumlah contoh aktual dari setiap kelas dalam data uji. Laporan ini menunjukkan bahwa distribusi data tidak seimbang, dengan kelas 1 memiliki 79 contoh, sementara kelas 4 hanya memiliki 14 contoh. Ketidakseimbangan ini bisa menjadi salah satu faktor yang menyebabkan rendahnya kinerja model pada beberapa kelas. Secara keseluruhan, accuracy dari model ini adalah 0.39 atau 39%, yang menunjukkan bahwa model belum bekerja dengan baik dalam mengklasifikasikan data ulasan ini. Macro average dan weighted average juga menunjukkan hasil yang rendah, dengan precision, recall, dan f1-score yang masing-masing berada di bawah 0.40. Ini mengindikasikan bahwa model mungkin perlu dioptimalkan lebih lanjut atau bahwa dataset ini memiliki karakteristik yang membuatnya sulit untuk diklasifikasikan dengan baik oleh model Naive Bayes. Penyebabnya bisa termasuk ketidakseimbangan data atau kompleksitas ulasan itu sendiri, yang mungkin memerlukan pendekatan yang lebih canggih untuk mencapai hasil yang lebih baik.

Persentase Ulasan:
Buruk: 87.50%
Tidak Buruk: 12.50%

Beberapa Hasil Prediksi dan Klasifikasi:			
	Content	Predicted Label	Classification
0	alfarro reyyza	1	Buruk
1	Martinah martinah	1	Buruk
2	Tania Martins	1	Buruk
3	Lizety Vidigal	1	Buruk
4	RAMA DANUR AL-HAQ	1	Buruk

Ulasan yang diklasifikasikan sebagai "Buruk":

	Content	Predicted Label
0	alfarro reyyza	1
1	Martinah martinah	1
2	Tania Martins	1
3	Lizety Vidigal	1
4	RAMA DANUR AL-HAQ	1
5	Muy Haya	1
7	Rizki 1108	1
8	Pengguna Google	1
9	Salman Alfarisi	1
10	Agung Wahyudi	1

Gambar 13. Persentase

Gambar 13 menunjukkan hasil dari analisis klasifikasi ulasan menggunakan model Naive Bayes, yang mengelompokkan ulasan ke dalam kategori "Buruk" dan "Tidak Buruk". Dari hasil analisis ini, kita melihat bahwa mayoritas ulasan yang dianalisis oleh model memiliki sentimen negatif, dengan 87.50% diklasifikasikan sebagai "Buruk" dan hanya 12.50% yang masuk dalam kategori "Tidak Buruk". Persentase ini mengindikasikan bahwa sebagian besar ulasan

yang diproses oleh model cenderung bersifat negatif. Bagian berikutnya dari hasil ini menampilkan beberapa contoh ulasan yang telah diprediksi dan diklasifikasikan oleh model. Pada bagian ini, setiap ulasan ditampilkan dengan teksnya (Content), label yang diprediksi oleh model (Predicted Label), dan hasil klasifikasinya (Classification), apakah ulasan tersebut tergolong "Buruk" atau "Tidak Buruk". Dari contoh yang ditampilkan, semua ulasan memiliki label prediksi 1 dan dikategorikan sebagai "Buruk", yang menunjukkan bahwa model konsisten dalam mengidentifikasi ulasan-ulasan ini sebagai sangat negatif. Selanjutnya, ditampilkan daftar ulasan spesifik yang dikelompokkan oleh model sebagai "Buruk". Ulasan-ulasan dari pengguna seperti "alfarro reyzya", "Martinah martinah", dan "Tania Martins" semuanya dikategorikan sebagai "Buruk" dengan label prediksi 1.

Ulasan yang diklasifikasikan sebagai 'Tidak Buruk':

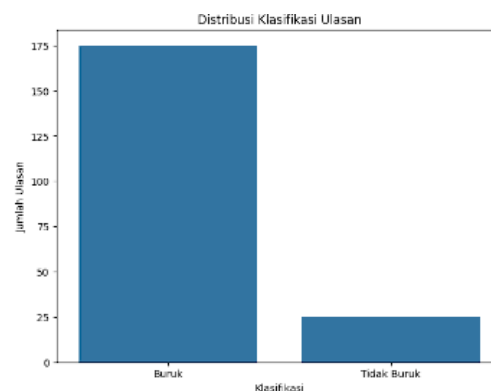
	Content	Predicted Label
6	Aqila and Elea chanel	5
12	Paya Costa	5
32	De Risma	5
37	Floriana Maria pirez	5
38	Ismail	5
44	Azzahra nurfadhillla	5
48	dewi nur	5
49	Alberto Santos	5
57	Sitait Ali	5
69	Aldo Lopes	5

Gambar 14. Tidak Buruk

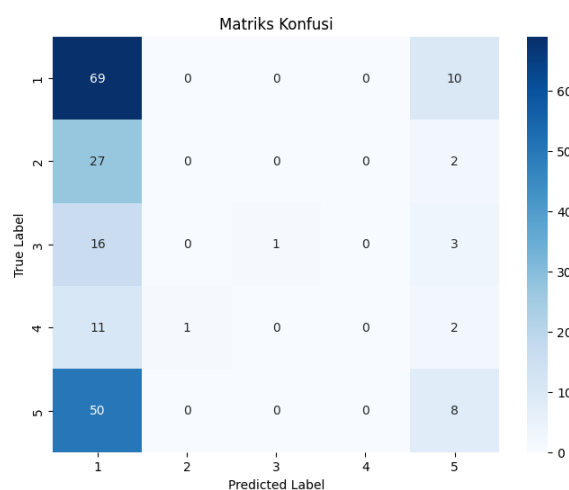
Gambar 14 melengkapi hasil klasifikasi ulasan yang dilakukan dengan model Naive Bayes, di mana ulasan-ulasan dikelompokkan ke dalam kategori "Tidak Buruk". Berbeda dengan kategori "Buruk" yang mendominasi dengan 87.50% dari total ulasan, kategori "Tidak Buruk" mencakup 12.50% dari ulasan yang dianalisis, menunjukkan bahwa hanya sebagian kecil ulasan yang dianggap memiliki sentimen positif atau netral. Bagian ini secara khusus menampilkan beberapa ulasan yang diklasifikasikan sebagai "Tidak Buruk". Ulasan-ulasan ini memiliki label prediksi 5, yang menunjukkan bahwa model Naive Bayes menilai ulasan-ulasan ini sebagai sangat positif atau memuaskan. Nama-nama seperti "Aqila and Elea chanel", "Paya Costa", dan "De Risma" muncul dalam daftar ini, menandakan bahwa ulasan dari pengguna-pengguna ini diprediksi memiliki sentimen positif yang kuat oleh model.

Gambar 15 menampilkan visualisasi distribusi klasifikasi ulasan yang dihasilkan oleh model Naive Bayes. Grafik batang ini menunjukkan perbandingan jumlah ulasan yang dikategorikan sebagai "Buruk" dan "Tidak Buruk" berdasarkan prediksi model. Dari grafik tersebut, terlihat dengan jelas bahwa sebagian besar ulasan diklasifikasikan sebagai "Buruk", dengan jumlah yang jauh lebih tinggi dibandingkan dengan ulasan yang masuk dalam kategori "Tidak Buruk". Hal ini konsisten dengan hasil yang telah ditampilkan

sebelumnya, di mana 87.50% dari total ulasan diklasifikasikan sebagai "Buruk", sementara hanya 12.50% yang dikategorikan sebagai "Tidak Buruk".



Gambar 15. Visualisasi Jumlah Ulasan



Gambar 16. Matriks

Gambar 16 menampilkan matriks konfusi hasil dari model Naive Bayes yang diterapkan pada data uji. Matriks konfusi ini memberikan wawasan mendalam tentang kinerja model dalam mengklasifikasikan ulasan ke berbagai kategori sentimen, serta mengidentifikasi kesalahan prediksi yang dilakukan oleh model. Pada matriks ini, sumbu vertikal menunjukkan True Label atau label asli dari ulasan, sedangkan sumbu horizontal menunjukkan Predicted Label atau label yang diprediksi oleh model. Diagonal utama matriks, yang berjalan dari kiri atas ke kanan bawah, mewakili jumlah prediksi yang benar—di mana label asli sesuai dengan label prediksi model. Misalnya, model berhasil memprediksi dengan benar 69 ulasan yang sebenarnya memiliki label 1. Namun, untuk label-label lain seperti 2, 3, dan 4, jumlah prediksi yang benar sangat sedikit, yang menunjukkan bahwa model mengalami kesulitan dalam membedakan antara variasi sentimen yang lebih halus. Selain itu, kesalahan prediksi yang signifikan terlihat di sel-sel di luar diagonal utama.

5. KESIMPULAN DAN SARAN

Bagian ini berisi ringkasan temuan utama penelitian yang menjawab tujuan penelitian. Kesimpulan disampaikan secara singkat dan jelas. Saran diberikan sebagai rekomendasi berdasarkan hasil penelitian

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa analisis menunjukkan sebagian besar ulasan pengguna dikategorikan sebagai "Buruk" oleh model Naive Bayes, dengan persentase mencapai 87,50%. Hal ini menunjukkan adanya ketidakpuasan yang dominan di antara pengguna terhadap aplikasi yang dianalisis. Model Naive Bayes juga menunjukkan keterbatasan dalam mengklasifikasikan ulasan, terutama untuk kategori yang lebih positif. Matriks konfusi memperlihatkan bahwa model cenderung memprediksi ulasan dengan label yang lebih rendah (negatif), yang mengindikasikan adanya bias dalam model. Kesalahan prediksi juga terlihat jelas, di mana banyak ulasan yang seharusnya memiliki sentimen positif atau netral justru diprediksi sebagai ulasan negatif. Ini menunjukkan bahwa model belum mampu menangkap variasi sentimen dengan baik dalam ulasan pengguna. Hasil ini mengindikasikan bahwa model Naive Bayes perlu dioptimalkan lebih lanjut, baik melalui penyesuaian data pelatihan, penggunaan metode preprocessing yang lebih baik, atau dengan mencoba algoritma lain yang lebih canggih untuk meningkatkan akurasi prediksi.

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, saran penelitian ini adalah sebagai berikut yaitu Pertama, disarankan untuk mencoba model pembelajaran mesin yang lebih kompleks, seperti Support Vector Machines (SVM) atau model berbasis neural networks, untuk melihat apakah mereka dapat memberikan hasil yang lebih baik dalam mengklasifikasikan sentimen ulasan. Model-model ini mungkin memiliki kemampuan yang lebih baik dalam menangani kompleksitas data dan memberikan akurasi prediksi yang lebih tinggi. Kedua, proses preprocessing data perlu diperbaiki dengan mempertimbangkan karakteristik khusus dari teks ulasan, seperti penggunaan bahasa informal dan slang, yang sering kali digunakan oleh pengguna dan dapat mempengaruhi hasil analisis sentimen. Dengan

memperhatikan aspek-aspek ini, diharapkan model dapat lebih baik menangkap variasi sentimen yang ada dalam ulasan pengguna.

DAFTAR PUSTAKA

- [1] D. N. Aliya and N. Yuliana, "Persepsi Publik di Indonesia mengenai Berita Hoaks di Media Sosial," *Sindoro Cendikia Pendidik.*, vol. 5, no. 1, pp. 61–70, 2024.
- [2] H. Febri, J. Siregar, R. Bangun, and K. Sukatman, "Strategi Efektif Memperkuat Hubungan Remaja Kristen dengan Tuhan Melalui Media Sosial di Era Digital," *J. Pendidik. Agama dan Teol.*, vol. 2, no. 3, pp. 129–141, 2024.
- [3] N. D. A. Amrita *et al.*, *Digital Marketing: Teori, Implementasi dan Masa Depan Digital Marketing*. PT. Green Pustaka Indonesia, 2024.
- [4] D. G. Nugroho, Y. H. Chrisnanto, and A. Wahana, "Analisis Sentimen Pada Jasa Ojek Online ... (Nugroho dkk.)," pp. 156–161, 2020.
- [5] A. N. Oktavia, M. Iqbal, R. W. Saputra, M. I. Zulfikar, and A. Saifudin, "Implementasi Metode Natural Language Processing Dalam Studi Analisis Semantik Dan Emosi Buzzer Pada Tweet Di Aplikasi X," *Bul. Ilm. Ilmu Komput. dan Multimed.*, vol. 2, no. 1, pp. 154–159, 2024.
- [6] Y. E. Rachmad *et al.*, *Integrasi metode kuantitatif dan kualitatif: Panduan praktis penelitian campuran*. PT. Green Pustaka Indonesia, 2024.
- [7] I. G. Hendrayana, D. Suprayitno, L. Judijanto, F. Kosadi, S. Y. Kusumastuti, and S. Sepriano, *E-Money: Panduan Lengkap Penggunaan dan Manfaat E-Money dalam Era Digital*. PT. Sonpedia Publishing Indonesia, 2024.
- [8] R. O. Firsta, "Pengaruh Penggunaan Media Sosial TikTok terhadap Minat Belajar Siswa pada Mata Pelajaran Pendidikan Agama Islam di SMP Negeri 2 Babadan Ponorogo," 2024, *IAIN Ponorogo*.
- [9] B. Ravinder, S. K. Seenii, V. S. Prabhu, P. Asha, S. P. Maniraj, and C. Srinivasan, "Web Data Mining with Organized Contents Using Naive Bayes Algorithm," in *2024 2nd International Conference on Computer, Communication and Control (IC4)*, IEEE, 2024, pp. 1–6.