

Analisis Sentimen Ulasan Platform Media Sosial Menggunakan Algoritma *Naive Bayes Classifier*

Sudin Saepudin^{[1]*}, Selviani Widiastuti^[2], Carti Irawan^[3]

Program Studi Sistem Informasi, Fakultas Teknik Komputer dan Desain^{[1], [2], [3]}

Universitas Nusa Putra

Sukabumi, Indonesia

sudin.saepudin@nusaputra.ic.id^[1], selviani.widiastuti_si19@nusaputra.ic.id^[2], carti@nusaputra.ic.id^[3]

Abstract— The Covid-19 pandemic has caused significant changes in people's lifestyles which are further strengthened by the rapid development of technology. This has resulted in increased use of the internet and accelerated dissemination of information through social media platforms. Not only for self-expression, social media can also be a means of communication, information, education, and even used as a marketing tool. Several social media platforms have recently been popular and widely used, the number of users is increasing from year to year, and each user can provide a rating review of the application. To find out public opinion on social media platforms, sentiment analysis will be carried out on several social media platform applications on the Google Play Store, namely Twitter, Instagram and Tiktok which will later be used as material for evaluating these applications. In this study, the dataset was taken based on ratings from user reviews on the Google Play Store using the NBC (Naive Bayes Classifier) method with the Python programming language. Based on testing of 1000 comment review data from each application, it was found that the majority gave positive sentiment (Twitter 57.2%, Instagram 74.1%, Tiktok 83.9%), and negative sentiment (Twitter 42.8%, Instagram 25.9%, Tiktok 16.1%) with an accuracy rate of 85.6% for the Twitter application, 83.6% for the Instagram application, and 84.8% for the Tiktok application.

Keywords— Sentiment Analysis, Google Play Store, NBC (Naive Bayes Classifier), Social Media Platform, Python.

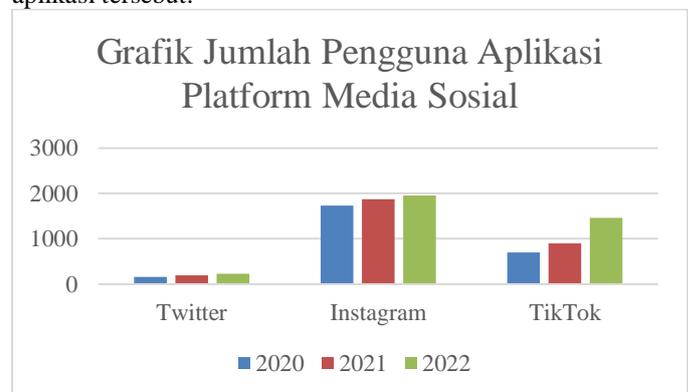
Abstrak— Pandemi Covid-19 telah menyebabkan perubahan yang signifikan dalam gaya hidup masyarakat yang semakin diperkuat oleh pesatnya perkembangan teknologi. Hal ini mengakibatkan peningkatan penggunaan internet dan percepatan penyebaran informasi melalui *platform* media sosial. Tak hanya untuk mengekspresikan diri, media sosial juga dapat menjadi alat komunikasi, informasi, edukasi, bahkan dijadikan alat pemasaran. Beberapa *platform* media sosial belakangan ini sedang populer dan marak digunakan, jumlah penggunanya semakin meningkat dari tahun ke tahun, dan setiap pengguna dapat memberikan ulasan penilaian terhadap aplikasi tersebut. Untuk mengetahui opini masyarakat terhadap *platform* media sosial maka akan dilakukan analisis sentimen pada beberapa aplikasi *platform* media sosial di Google Play Store yaitu Twitter, Instagram, dan Tiktok yang nantinya akan digunakan sebagai bahan evaluasi terhadap aplikasi tersebut. Dalam penelitian ini dataset diambil berdasarkan rating dari review pengguna di Google Play Store menggunakan metode NBC (*Naive Bayes Classifier*) dengan bahasa pemrograman Python. Berdasarkan pengujian dari 1000 data ulasan komentar dari masing-masing

aplikasi, didapatkan sebagian besar memberi sentimen positif (Twitter 57,2%, Instagram 74,1%, Tiktok 83,9%), dan sentimen negatif (Twitter 42,8%, Instagram 25,9%, Tiktok 16,1%) dengan nilai tingkat akurasi 85,6% untuk aplikasi Twitter, 83,6% untuk aplikasi Instagram, dan 84,8% untuk aplikasi Tiktok.

Kata Kunci— Analisis Sentimen, Google Play Store, NBC (*Naive Bayes Classifier*), Platform Media Sosial, Python.

I. PENDAHULUAN

Pandemi Covid-19 telah menyebabkan perubahan yang signifikan dalam gaya hidup masyarakat yang semakin diperkuat oleh pesatnya perkembangan teknologi. Hal ini mengakibatkan peningkatan penggunaan internet dan percepatan penyebaran informasi melalui *platform* media sosial. Pada saat ini *platform* media sosial sedang marak digunakan untuk media komunikasi, informasi, edukasi, bahkan dijadikan media pemasaran. Ada beberapa *platform* media sosial yang belakangan ini sedang populer dengan fitur serta kegunaan yang hampir sama yaitu Tiktok, Instagram, dan Twitter. Jumlah pengguna aplikasi tersebut semakin meningkat dari tahun ke tahun semenjak adanya pandemi. Setelah melakukan studi literatur dari beberapa sumber, peneliti menemukan data tiga tahun kebelakang jumlah pengguna dari aplikasi Twitter, Instagram, dan Tiktok di website Databoks – Katadata. Berikut ini grafik jumlah pengguna dari ketiga aplikasi tersebut.



Gambar 1. Grafik Jumlah Pengguna Aplikasi Platform Media Sosial

Dari Gambar 1, dapat dilihat perbandingan jumlah

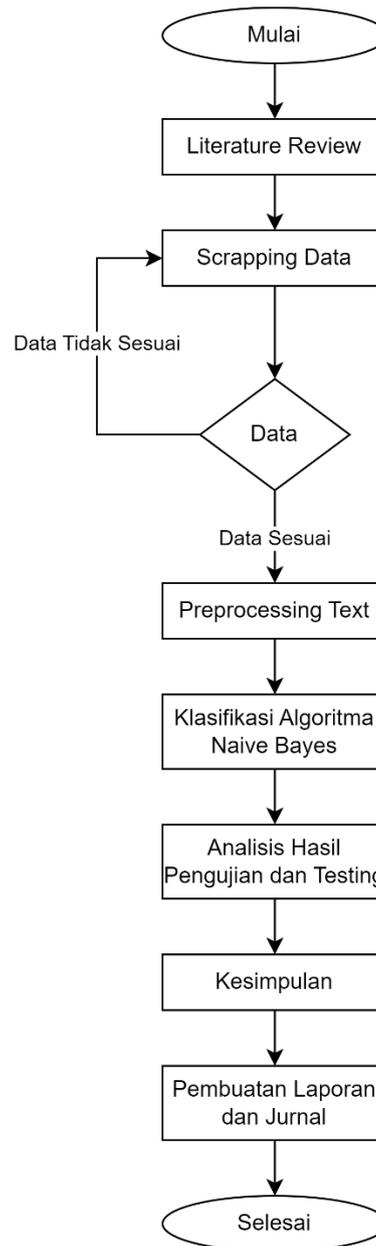
pengguna dari masing-masing aplikasi *platform media sosial* dari tahun 2020-2022, dimana aplikasi Twitter pada tahun 2020 memiliki sebanyak 166 juta pengguna, pada tahun 2021 memiliki sebanyak 199 juta pengguna, dan pada tahun 2022 memiliki sebanyak 229 juta pengguna, selanjutnya aplikasi Instagram pada tahun 2020 memiliki sebanyak 1,734 miliar pengguna, pada tahun 2021 memiliki sebanyak 1,878 miliar pengguna, dan pada tahun 2022 memiliki sebanyak 1,961 miliar pengguna, terakhir yaitu aplikasi Tiktok pada tahun 2020 memiliki sebanyak 700 juta pengguna, pada tahun 2021 memiliki sebanyak 902 juta pengguna, dan pada tahun 2022 memiliki sebanyak 1,460 miliar pengguna.

Dari latar belakang masalah yang ada maka dilakukan analisis sentimen pada penelitian ini untuk mengetahui opini masyarakat tentang platform media sosial di Google Play Store [1]. Beberapa penelitian terkait dengan Analisis Sentimen telah dilakukan sebelumnya, diantaranya penelitian yang dilakukan oleh Sola Fide, dkk tahun 2021. Peneliti tersebut melakukan analisis sentimen pada aplikasi Tiktok dengan menggunakan data sebanyak 3200 ulasan dari September 2020 hingga Februari 2021 di Google Play, pada aplikasi Tiktok ada lebih banyak ulasan positif daripada ulasan negatif, dengan ulasan positif sebanyak 1741 (54.41%), dan ulasan negatif sebanyak 1459 (45.59%). Dengan menerapkan metode *Support Vector Machine* (SVM) dan membagi data *training* dan data *testing* dengan perbandingan 80:20, klasifikasi sentimen dari hasil *sentiment scoring review* aplikasi TikToc di Google Play mencapai tingkat akurasi, yaitu 81,24% [2]. Penelitian selanjutnya yang dilakukan oleh Sifa Melina Salsabila, dkk tahun 2022. Peneliti mengambil data dari Google Play Store. Algoritma *Naive Bayes Classifier* digunakan dalam menganalisis ulasan pengguna aplikasi Tokopedia. Setelah dilakukan pengujian oleh rapid evaluator pada 143 data pengujian, diperoleh hasil akurasi sebesar 95,10% [3]. Penelitian selanjutnya oleh Cholid Fadilah Hasri, dkk tahun 2022. Peneliti menggunakan twitter untuk mengumpulkan data pada penelitian ini. Penelitian ini menggunakan metode *Naive Bayes Classifier* (NBC) sebagai algoritma utama untuk menganalisis sentimen masyarakat terhadap dampak virus corona. Hasil penerapan metode Naive Bayes menunjukkan nilai akurasi sebesar 81,07%. Selain itu, untuk memastikan keakuratan hasil penelitian, dilakukan pengujian menggunakan metode *Support Vector Machine* yang menghasilkan nilai akurasi sebesar 79,96% [4].

Meskipun telah banyak penelitian yang dilakukan dalam bidang analisis sentimen, namun masih sedikit penelitian yang secara komprehensif menganalisis sentimen pengguna terhadap *platform media sosial* Twitter, Instagram, dan TikTok. Oleh karena itu, penelitian ini bertujuan untuk mengisi kesenjangan penelitian tersebut dengan melakukan analisis sentimen terhadap ulasan pengguna pada ketiga platform media sosial tersebut menggunakan algoritma *Naive Bayes Classifier*, *tools* Google Colab, dan bahasa pemrograman Python.

II. METODE PENELITIAN

Penelitian ini melalui beberapa tahapan dimulai dari *web scraping*, *preprocessing text*, klasifikasi algoritma NBC, analisis hasil pengujian dan *testing*.



Gambar 2. Sistematika Penelitian

A. Web Scraping

Web scraping adalah cara untuk memperoleh informasi dari sebuah situs web secara otomatis tanpa perlu melakukan penyalinan secara manual [5]. *Web scraping* dapat didefinisikan juga sebagai metode untuk melakukan ekstraksi informasi tertentu dari situs web yang dituju, kemudian dapat dilakukan analisis dimanfaatkan untuk berbagai tujuan [6].

B. Preprocessing Text

Pre-processing merupakan tahapan awal pengolahan data menjadi analisa sentimen yang bertujuan untuk

menghilangkan kata atau kalimat yang tidak perlu dan menyamakan bentuk kata untuk mengurangi volume pada kosakata [7]. Langkah-langkah pada *pre-processing* yaitu sebagai berikut :

1) *Case Folding*

Case folding merujuk pada proses mengubah semua huruf pada dokumen atau kalimat menjadi huruf kecil [8]. Proses tersebut dilakukan untuk mempermudah dan mempercepat pembacaan oleh mesin pada korpus yang dimaksud [9].

2) *Tokenizing*

Tokenisasi adalah proses memisahkan rangkaian karakter berdasarkan spasi, serta memungkinkan penghapusan karakter tertentu seperti tanda baca pada waktu yang bersamaan [10]. Tujuan dari dilakukannya tokenisasi ini adalah agar dalam proses selanjutnya tidak memproses berdasarkan kalimat lagi tetapi memproses kata per kata [9].

3) *Cleansing*

Cleansing adalah proses pembersihan data teks dengan cara menghapus data yang tidak konsisten atau tidak diperlukan. *Cleansing* dilakukan guna menganalisa kualitas data yang memiliki format data atau file yang salah guna menghasilkan data berkualitas tinggi [11].

4) *Normalization*

Normalisasi adalah proses perubahan kata tidak baku pada data komentar yang telah dikumpulkan menjadi kata-kata standar yang sesuai dengan Kamus Besar Bahasa Indonesia (KBBI). Tujuannya adalah untuk memastikan kata-kata yang diproses memiliki arti yang sesuai dengan KBBI [9].

5) *Stemming*

Kata-kata dari data komentar yang telah melewati beberapa proses sebelumnya, kemudian dilakukan pengecekan kembali yang merupakan tahapan akhir dalam *preprocessing* [10]. Proses tersebut disebut *stemming* yang tujuannya menghilangkan semua imbuhan yang ada pada suatu kata baik yang berposisi di awal, akhir, atau kombinasi dari keduanya [9].

melakukan perhitungan [13].

Metode klasifikasi teks *Naïve Bayes Classifier* terdiri dari dua tahap, yaitu tahap pelatihan dan tahap klasifikasi. Pada tahap pelatihan, dilakukan proses terhadap sampel data yang diharapkan dapat merepresentasikan data tersebut. Sedangkan pada tahap klasifikasi, dilakukan penentuan probabilitas prior bagi setiap kategori berdasarkan sampel data pada tahap pelatihan [15].

III. HASIL DAN PEMBAHASAN

Bagian ini akan menjelaskan tentang hasil dari penelitian yang telah dilakukan, dimulai dari proses pengumpulan data, klasifikasi *naïve bayes*, hasil klasifikasi (*Accuracy*, *Precision*, *Recall* dan *f1-score*), dan visualisasi data.

A. Pengumpulan Data

Pengumpulan data menggunakan metode *scraping* dengan data yang diperoleh sebanyak 1000 ulasan komentar dari masing-masing aplikasi pada website Google Play Store.

TABEL 1. PROSES *PRE-PROCESSING* APLIKASI TWITTER

User Name	Content	Score	At
jayani tinenta	Biasa	2	2023-02-23 08:22:11
Natan Markus	Akun yang tidak pernah posting apa-apa pun tet...	1	2023-02-23 08:20:20
Angga Saputra	Ga bisa buat mutar video. Tiap hari harus unin...	1	2023-02-23 07:56:01
Arsikin Sugeng	Saya belum bisa login Twitter	4	2023-02-23 07:50:06
achmad Naufal	#Quoteshariini Jangan pantang menyerah	1	2023-02-23 07:47:45

C. Klasifikasi Algoritma NBC

Naïve Bayes Classifier merupakan sebuah pengklasifikasi yang paling sederhana dan paling sering digunakan. *Naïve Bayes* adalah machine learning yang menggunakan probabilitas dengan konsep pendekatan *Bayes* sebagai perhitungannya [12]. Perhitungan probabilitas teorema Bayes pada *Naïve Bayes Classifier* mengasumsikan bahwa nilai atribut dari suatu kelas tidak memiliki keterkaitan dengan keberadaan atribut kelas lain, meskipun kenyataannya hampir setiap atribut memiliki ketergantungan satu sama lain [13]. Dapat disimpulkan, *Naïve Bayes* merupakan metode untuk klasifikasi teks dengan kecepatan pemrosesan yang tinggi untuk data yang besar [14]. Dengan asumsi yang dibuat seperti itu, pada akhirnya membuat *Naïve Bayes* mudah dalam

TABEL 2. PROSES *PRE-PROCESSING* APLIKASI INSTAGRAM

User Name	Content	Score	At
siggit semito adi	Ok.	5	2023-02-23 07:05:33
Muhammad Khoirulmam	Apk aneh engak bisa buat login	1	2023-02-23 07:05:25
Mohamad Wahyudi	sip pokok wkwkwkwk anjay follow ig ku lurr ☺	5	2023-02-23 07:03:30
123 GO!!! INDO	Saya sangat senang dengan akun instagram sebag...	5	2023-02-23 07:03:23
Naryo Naeryl123	Love Instagram	5	2023-02-23 07:00:22

TABEL 3. PROSES *PRE-PROCESSING* APLIKASI TIKTOK

User Name	Content	Score	At
KTUK Budayasa	Aplikasi ini sangat bagus	1	2023-02-23 07:15:02
Ridho Saputra	Bagus keren kayak babi	5	2023-02-23 07:13:50
ALVARO NET	Ok	5	2023-02-23 07:12:59
Ilda Rasyida	apps nya bagus banget ga nyesel deh	5	2023-02-23 07:12:54
Icha Chelsea	Apk ini bagus	5	2023-02-23 07:12:18

Pada Tabel 4, Tabel 5, dan Tabel 6, terdapat proses *pre-processing* data, yaitu data mentah hasil *scraping* akan melalui tahap *case folding* (data diubah menjadi huruf kecil) dan juga dilakukan *cleaning* (menghilangkan karakter, url, hashtag dan simbol) hingga mendapatkan hasil komentar review. Selanjutnya dilakukan pemotongan kata pada kalimat komentar yang dipisahkan oleh tanda koma (,) pada tahap *tokenizing*. Lalu akan dilakukan proses *filtering* untuk menghilangkan kata-kata tidak bermanfaat. Tahapan akhir *pre-processing* yaitu *stemming* dimana akan memastikan data yang mengandung kata imbuhan akan dihilangkan agar mendapatkan kumpulan data yang berisi kata-kata dasar.

TABEL 4. HASIL *PRE-PROCESSING* APLIKASI TWITTER

Case Folding dan Cleaning	Tokenizing	Filtering	Stemming
Biasa	[biasa]	[]	NaN
akun yang tidak pernah posting apa-apa pun tet...	[akun, yang, tidak, pernah, posting, apa-apa, ...	[akun, posting, apa-apa, ditangguhkan]	akun posting apa Tangguh
ga bisa buat mutar video tiap hari harus uninst...	[ga, bisa, buat, mutar, video, tiap, hari, har...	[ga, mutar, video, uninstall, aplikasi, downlo...	ga mutar video uninstall aplikasi download bia...
saya belum bisa login twitter	[saya, belum, bisa, login, twitter]	[login, twitter]	login twitter
#quoteshariini jangan pantang menyerah	[#quoteshariini, jangan, pantang, menyerah]	[#quoteshariini, pantang, menyerah]	quoteshariini pantang serah

TABEL 5. HASIL *PRE-PROCESSING* APLIKASI INSTAGRAM

Case Folding dan Cleaning	Tokenizing	Filtering	Stemming
Ok	[ok]	[ok]	Ok
apk aneh engak bisa buat login	[apk, aneh, engak, bisa, buat, login]	[apk, aneh, engak, login]	apk aneh engak login
sip pokok wkwkwkwk anjay follow ig ku lurr ☺	[sip, pokok, wkwkwkwk, anjay, follow, ig, ku, ...	[sip, pokok, wkwkwkwk, anjay, follow, ig, ku, ...	sip pokok wkwkwkwk anjay follow ig ku lurr
saya sangat senang dengan akun instagram sebag...	[saya, sangat, senang, dengan, akun, instagram...]	[senang, akun, instagram, media, menyalurkan, ...	senang akun instagram media salur inspirasi ma...
love Instagram	[love, instagram]	[love, instagram]	love instagram

TABEL 6. HASIL *PRE-PROCESSING* APLIKASI TIKTOK

Case Folding dan Cleaning	Tokenizing	Filtering	Stemming
aplikasi ini sangat bagus	[aplikasi, ini, sangat, bagus]	[aplikasi, bagus]	aplikasi bagus
bagus keren kayak babi	[bagus, keren, kayak, babi]	[bagus, keren, kayak, babi]	bagus keren kayak babi
Ok	[ok]	[ok]	Ok
apps nya bagus banget ga nyesel deh	[apps, nya, bagus, banget, ga, nyesel, deh]	[apps, nya, bagus, banget, ga, nyesel, deh]	apps nya bagus banget ga nyesel deh
apk ini bagus	[apk, ini, bagus]	[apk, bagus]	apk bagus

B. Klasifikasi Naïve Bayes

Proses klasifikasi dilakukan menggunakan *Multinomial Naïve Bayes* dengan metode *splitting* (80:20), 80% data akan digunakan sebagai training data, dan 20% data akan digunakan sebagai *testing data*. Setelah melalui tahap pengujian dimana yang awalnya dari masing-masing aplikasi diambil 1000 data, beberapa komentar dihilangkan karena nilai *rating* adalah bias sehingga data yang akan digunakan untuk aplikasi Twitter sebanyak 940, aplikasi Instagram sebanyak 943, dan aplikasi Tiktok sebanyak 955. Data yang digunakan merupakan *random* dari dataset yang telah dibersihkan.

TABEL 7. MATRIKS PREDIKSI APLIKASI TWITTER

Prediksi	Aktual	
	Negatif	Positif
Negatif	56	18
Positif	9	105

TABEL 8. MATRIKS PREDIKSI APLIKASI INSTAGRAM

Prediksi	Aktual	
	Negatif	Positif
Negatif	25	30
Positif	1	133

TABEL 9. MATRIKS PREDIKSI APLIKASI TIKTOK

Prediksi	Aktual	
	Negatif	Positif
Negatif	2	29
Positif	0	160

Confusion matrix :

- TN (*True Negative*): Model memprediksi data terdapat pada kelas Negatif dan yang sebenarnya data memang terdapat pada kelas Negatif.
- TP (*True Positive*): Model memprediksi data terdapat pada kelas Positif dan yang sebenarnya data memang terdapat di kelas Positif.
- FN (*False Negative*): Model memprediksi data berada di kelas Negatif, tetapi data sebenarnya ada di kelas Positif.

- FP (*False Positive*): Model memprediksi data berada di kelas Positif, tetapi data sebenarnya ada di kelas Negatif.

Dari hasil pengujian menggunakan *multinomial Naive Bayes* yang dilakukan dengan ulasan komentar aplikasi Twitter sebanyak 188 data, aplikasi Instagram sebanyak 189 data, dan aplikasi Tiktok sebanyak 191 data (20% dari dataset secara random). Dari Tabel 7, Tabel 8, dan Tabel 9 didapatkan hasil *confusion matrix*:

- TN aplikasi Twitter sebanyak 56, aplikasi Instagram sebanyak 25, dan aplikasi Tiktok sebanyak 2,
- FP aplikasi Twitter sebanyak 18, aplikasi Instagram sebanyak 30, dan aplikasi Tiktok sebanyak 29,
- TP aplikasi Twitter sebanyak 105, aplikasi Instagram sebanyak 133, dan aplikasi Tiktok sebanyak 160,
- FN aplikasi Twitter sebanyak 9, aplikasi Instagram sebanyak 1, dan aplikasi Tiktok sebanyak 0.

C. Hasil Klasifikasi (*Accuracy, Precision, Recall dan f1-score*)

Selanjutnya, dilakukan proses *cross-validation* menggunakan algoritma *Naïve Bayes Classifier* (NBC). Data akan melewati tahap pengujian (*testing*) dan pelatihan (*training*), kemudian dilakukan pemodelan dengan menggunakan algoritma NBC untuk mendapatkan nilai akurasi (*accuracy*), presisi (*precision*), *recall*, dan skor F1 (*f1 score*) berdasarkan performa algoritma *Machine Learning* (ML). Evaluasi ini didasarkan pada *Confusion Matrix*. Berikut adalah hasil dari klasifikasi menggunakan algoritma Naive Bayes dengan pembobotan TF-IDF:

```
MultinomialNB Accuracy: 0.8563829787234043
MultinomialNB Precision: 0.8615384615384616
MultinomialNB Reccal: 0.7567567567567568
MultinomialNB f1_score: 0.8057553956834532
Confusion Matrix:
[[ 56 18]
 [ 9 105]]
```

Gambar 3. Hasil Klasifikasi Algoritma *Multinomial Naive Bayes* Pada Aplikasi Twitter

```
MultinomialNB Accuracy: 0.8359788359788359
MultinomialNB Precision: 0.9615384615384616
MultinomialNB Reccal: 0.45454545454545453
MultinomialNB f1_score: 0.617283950617284
Confusion Matrix:
[[ 25 30]
 [ 1 133]]
```

Gambar 4. Hasil Klasifikasi Algoritma *Multinomial Naive Bayes* Pada Aplikasi Instagram

```
MultinomialNB Accuracy: 0.8481675392670157
MultinomialNB Precision: 1.0
MultinomialNB Reccal: 0.06451612903225806
MultinomialNB f1_score: 0.12121212121212122
Confusion Matrix:
[[ 2 29]
 [ 0 160]]
```

Gambar 5. Hasil Klasifikasi Algoritma *Multinomial Naive Bayes* Pada Aplikasi Tiktok

Berikut ini merupakan penjelasan dari Gambar 3, Gambar 4, dan Gambar 5, nilai Precision (Perbandingan data antara TP “*True Positive*” dengan banyak data yang diprediksi Positif) dihasilkan menggunakan algoritma *Multinomial Naive Bayes* didapatkan :

- Aplikasi Twitter sebesar 0.86153
- Aplikasi Instagram sebesar 0.96153
- Aplikasi Tiktok sebesar 1.0

Nilai *Recall* (Perbandingan banyak data antara TP “*True Positive*” dengan banyak data yang sebenarnya positif) yang didapatkan :

- Aplikasi Twitter sebesar 0.75675
- Aplikasi Instagram sebesar 0.45454
- Aplikasi Tiktok sebesar 0.06451

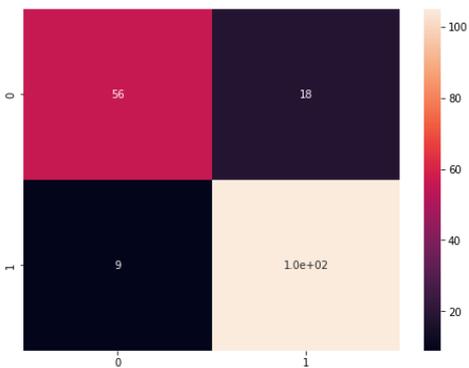
F1-Score (Nilai *harmonic* antara *Precision* dengan *Recall*) yang didapatkan :

- Aplikasi Twitter sebesar 0.80575
- Aplikasi Instagram sebesar 0.61728
- Aplikasi Tiktok sebesar 0.12121

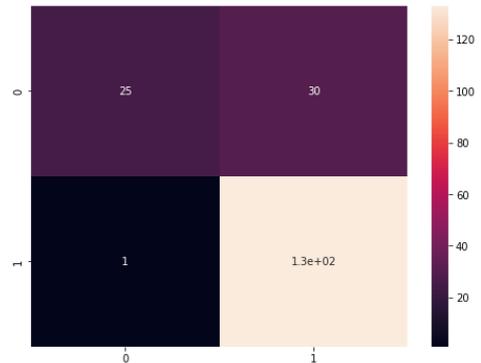
Dengan keakuratan algoritma yang digunakan sebesar 0.85638 untuk aplikasi Twitter, 0.83597 untuk aplikasi Instagram, dan 0.84816 untuk aplikasi Tiktok. Dapat dikatakan bahwa model klasifikasi memiliki nilai yang sangat baik.

D. Visualisasi Data

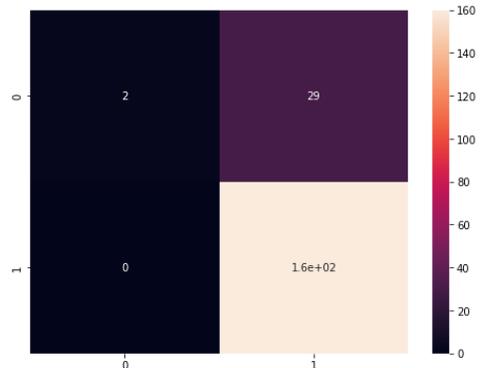
Dari hasil pengklasifikasian menggunakan *Multinomial Naive Bayes* didapatkan *Confusion matrix* yang divisualisasikan pada Gambar 6, Gambar 7, dan Gambar 8.



Gambar 6. *Confusion Matrix* Aplikasi Twitter



Gambar 7. *Confusion Matrix* Aplikasi Instagram



Gambar 8. *Confusion Matrix* Aplikasi Tiktok

Adapun kata-kata yang sering muncul dalam teks ulasan akan diperbaiki dan divisualisasikan menggunakan *wordcloud*. Visualisasi kata yang sering muncul ditunjukkan pada Gambar 9, Gambar 10, dan Gambar 11.



Gambar 9. *Wordcloud* Ulasan Aplikasi Twitter



Gambar 10. *Wordcloud* Ulasan Aplikasi Instagram

Hal ini menunjukkan bahwa pengguna aplikasi Instagram dan TikTok cenderung memberikan ulasan dengan sentimen positif lebih banyak dibandingkan dengan ulasan negatif. Sedangkan pengguna aplikasi Twitter memiliki kecenderungan ulasan yang lebih seimbang antara sentimen positif dan negatif. Analisis hasil ini memberikan pemahaman yang lebih dalam tentang pandangan dan tanggapan pengguna terhadap ketiga platform media sosial tersebut. Dalam konteks ini, aplikasi TikTok terlihat memiliki tingkat sentimen positif yang paling tinggi, diikuti oleh Instagram dan Twitter.

IV. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis sentimen yang telah dilakukan pada 1000 ulasan pengguna dari masing-masing aplikasi Twitter, Instagram, dan TikTok, menggunakan algoritma *Naive Bayes Classifier*, diperoleh kesimpulan sebagai berikut: Sebagian besar pengguna memberikan sentimen positif terhadap aplikasi Twitter, Instagram, dan TikTok. Persentase sentimen positif tertinggi ditemukan pada aplikasi TikTok (83,9%), diikuti oleh Instagram (74,1%) dan Twitter (57,2%). Dalam visualisasi data menggunakan wordcloud, kata-kata yang sering muncul dalam teks ulasan menunjukkan bahwa aplikasi Twitter dan Tiktok memiliki keunggulan dalam hal video, sedangkan aplikasi Instagram memiliki keunggulan dalam fitur *story*. Model prediksi yang digunakan dalam penelitian ini mampu menghasilkan prediksi dengan tingkat akurasi yang tinggi, yaitu 85,6% untuk Twitter, 83,6% untuk Instagram, dan 84,8% untuk Tiktok.

REFERENCES

- [1] S. Fide, Suparti, and Sudarno, "Analisis Sentimen Ulasan Aplikasi Tiktok Di Google Play Menggunakan Metode Support Vector Machine (SVM) Dan Asosiasi," *J. Gaussian*, vol. 10, no. 3, pp. 346–358, 2021.
- [2] S. M. Salsabila, A. A. Murtopo, and N. Fadhilah, "Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naive Bayes Classifier," *J. Minfo Polgan*, vol. 11, no. 02, 2022, [Online]. Available: www.tokopedia.com
- [3] C. F. Hasri and D. Alita, "Penerapan Metode Naive Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 3, no. 2, pp. 145–160, 2022, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/informatika>
- [4] M. F. Al-shufi and A. Erfina, "Sentimen Analisis Mengenai Aplikasi Streaming Film Menggunakan Algoritma Support Vector Machine Di Play Store," *J. SISMATIK*, vol. 01, no. 01, 2021, Accessed: Apr. 14, 2023. [Online]. Available: <https://sismatik.nusaputra.ac.id/index.php/sisimatik/article/view/22>
- [5] N. Ramadhani and N. Fajarianto, "Sistem Informasi Evaluasi Perkuliahan dengan Sentimen Analisis Menggunakan Naive Bayes dan Smoothing Laplace," *J. Sist. Inf. BISNIS*, vol. 10, no. 2, pp. 228–234, Dec. 2020, doi: 10.21456/vol10iss2pp228-234.
- [6] M. T. Nitami and H. Februariyaniti, "Analisis Sentimen Ulasan Ekspedisi J&T Express Menggunakan Algoritma Naive Bayes," *Jurna Manaj. Inform. Sist. Inf.*, vol. 5, no. 1, 2022.
- [7] R. Wahyudi and G. Kusumawardhana, "Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine," *J. Inform.*, vol. 8, no. 2, 2021, [Online]. Available: <http://ejournal.bsi.ac.id/ejournal/index.php/ji>
- [8] C. Villavicencio, J. J. Macrohon, X. A. Inbaraj, J. H. Jeng, and J. G. Hsieh, "Twitter sentiment analysis towards covid-19 vaccines in the Philippines using naïve bayes," *Inf.*, vol. 12, no. 5, May 2021, doi: 10.3390/info12050204.
- [9] M. Wongkar and A. Angdresy, "Sentiment analysis using Naive Bayes Algorithm of the data crawler: Twitter," in *2019 Fourth International Conference on Informatics and Computing (ICIC)*, IEEE, 2019, pp. 1–5.
- [10] S. Satriajati, S. B. Panuntun, and S. Pramana, "Implementasi Web Scraping Dalam Pengumpulan Berita Kriminal Pada Masa Pandemi Covid-19 Studi Kasus: Situs Berita detik.com," *Semin. Nas. Off. Stat.*, no. 01, 2020.
- [11] M. Djufri, "Penerapan Teknik Web Scraping Untuk Penggalan Potensi Pajak (Studi Kasus pada Online Market Place Tokopedia, Shopee dan Bukalapak)," *J. BPPK*, vol. 13, no. 02, pp. 65–75, 2020.
- [12] D. Normawati and S. A. Prayogi, "Implementasi Naive Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021.
- [13] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *J. TEKNO KOMPAK*, vol. 15, no. 1, 2021.
- [14] R. P. Sidiq, B. A. Dermawan, and Y. Umaidah, "Sentimen Analisis Komentar Toxic pada Grup Facebook Game Online Menggunakan Klasifikasi Naive Bayes," *J. Inform. Univ. Pamulang*, vol. 5, no. 3, p. 356, Sep. 2020, doi: 10.32493/informatika.v5i3.6571.
- [15] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," *JEPIN*, vol. 4, no. 2, pp. 17–29, 2018, [Online]. Available: www.femaledaily.com
- [16] W. Irmayani, "Visualisasi Data Pada Data Mining Menggunakan Metode Klasifikasi Naive Bayes," *J. Khatulistiwa Inform.*, vol. 9, no. 1, 2021, [Online]. Available: www.bsi.ac.id
- [17] N. Purwati, R. Nurlistiani, and O. Devinsen, "Data Mining Dengan Algoritma Neural Network Dan Visualisasi Data Untuk Prediksi Kelulusan Mahasiswa," *J. Inform.*, vol. 20, no. 2, pp. 156–163, Dec. 2020, doi: 10.30873/ji.v20i2.2273.
- [18] Z. U. Siregar, R. R. A. Siregar, and R. Arianto, "Klasifikasi Sentiment Analysis Pada Komentar Peserta Diklat Menggunakan Metode K-Nearest Neighbor," *J. KILAT*, vol. 8, no. 1, 2019.
- [19] W. Yulita, E. D. Nugroho, and M. H. Algifari, "Sentiment Analysis on Public Opinion About the Covid-19 Vaccine Using the Naive Bayes Classifier Algorithm." *Jdmsi*, 2021.
- [20] A. E. Sari, S. Widowati, and K. M. Lhaksmana, "Klasifikasi Ulasan Pengguna Aplikasi Mandiri Online di Google Play Store dengan Menggunakan Metode Information Gain dan Naive Bayes Classifier," *e-Proceeding Eng.*, vol. 6, no. 2, 2019.
- [21] R. Puspita and A. Widodo, "Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS," *J. Inform. Univ. Pamulang*, vol. 5, no. 4, p. 646, 2021.
- [22] D. Rusdian and D. Rosiyadi, "Analisa Sentimen Terhadap Tokoh Publik Menggunakan Metode Naive Bayes Classifier Dan Support Vector Machine," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 4, no. 2, pp. 2502–7131, 2019.
- [23] G. L. A. K. Putra and G. P. P. A. Yasa, "Komik Sebagai Sarana Komunikasi Promosi Dalam Media Sosial," *J. NAWALA Vis.*, vol. 1, no. 1, 2019, [Online]. Available: <https://jurnal.std-bali.ac.id/index.php/nawalavisual>
- [24] A. E. Y. Wibawa, "Implementasi Platform Digital Sebagai Media Pembelajaran Daring Di MI Muhammadiyah PK Kartasura Pada Masa Pandemi Covid-19," *Berajah J.*, vol. 01, no. 02, 2021, [Online]. Available: www.belajar.id
- [25] D. S. Utami and A. Erfina, "Analisis Sentimen Objek Wisata Bali Di Google Maps Menggunakan Algoritma Naive Bayes," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 6, no. 1, pp. 418–427, 2022.