

Analisis Sentimen Emosi Publik Terhadap Kebijakan iPhone 16 Menggunakan Algoritma *Naive Bayes*

Feisya Kartika Ilmi¹, Syamsa Shahira Julyinda², Eva Mahdyta Kiswana³, M. Yoka Fathoni⁴

*Fakultas Rekayasa Industri, Universitas Telkom Purwokerto
Jl. DI Panjaitan No.128, Karangreja, Purwokerto Kidul, Kec. Purwokerto Sel., Kabupaten Banyumas, Jawa Tengah, Indonesia*

¹21103076@ittelkom-pwt.ac.id

²21103078@ittelkom-pwt.ac.id

³21103059@ittelkom-pwt.ac.id

⁴myokafathoni@telkomuniversity.ac.id

Dikirim pada 22-11-2024, Direvisi pada 27-11-2024, Diterima pada 04-12-2024

Abstrak

Kebijakan pemerintah Indonesia yang melarang penjualan iPhone 16 akibat tidak terpenuhinya Persyaratan Tingkat Komponen Dalam Negeri (TKDN) memicu berbagai tanggapan di masyarakat. Media sosial, khususnya Twitter, menjadi platform utama bagi publik untuk menyuarakan opini mereka terhadap kebijakan ini. Tujuan dari penelitian ini adalah untuk menganalisis sentimen pengguna, media sosial terhadap kebijakan pemerintah Indonesia yang melarang Penjualan iPhone 16 akibat tidak terpenuhinya Persyaratan Tingkat Komponen Dalam Negeri (TKDN). Data dikumpulkan dari twitter dengan kata kunci relevan dan diproses menggunakan teknik *preprocessing* seperti *translate data*, *case folding* dan *tokenisasi*. Metode *Term Frequency Inverse Document Frequency* (TF-IDF) pada penelitian ini digunakan untuk ekstraksi fitur, sementara algoritma *Naive Bayes* diterapkan untuk klasifikasi sentimen menjadi positif, negatif, dan netral. Hasil evaluasi menunjukkan bahwa model memiliki akurasi sebesar 0,718182, dengan *precision* 0,749023, *recall* 0,718182, dan F1-score 0,714721. Temuan ini menggambarkan persepsi masyarakat yang beragam terhadap kebijakan pemerintah, dengan kontribusi signifikan dalam memahami opini publik melalui pendekatan kuantitatif berbasis data besar. Penelitian juga membuka peluang pengembangan model yang lebih terbuka untuk mengatasi kompleksitas data sosial media.

Kata Kunci: *Analisis Sentimen, Naive Bayes, Pemerintah, iPhone 16*

Ini adalah artikel akses terbuka di bawah lisensi [CC BY-SA](#).



Penulis Koresponden:

Feisya Kartika Ilmi

Jl. DI Panjaitan No.128, Karangreja, Purwokerto Kidul, Kec. Purwokerto Sel., Kabupaten Banyumas, Jawa Tengah, Indonesia

Email: 21103076@ittelkom-pwt.ac.id

I. PENDAHULUAN

Platform media X merupakan salah satu platform media sosial yang sangat digemari di Indonesia. Penggunaannya datang dari berbagai latar belakang dan bidang keahlian, sehingga diskusi yang terjadi di Twitter sering kali mencerminkan pandangan yang luas dan beragam terhadap berbagai isu. Dengan karakteristik ini, Twitter sering digunakan sebagai sarana untuk mengekspresikan opini dan kritik, termasuk yang berkaitan dengan kebijakan pemerintah. Salah satu topik yang baru-baru ini menjadi sorotan adalah keputusan pemerintah Indonesia untuk melarang penjualan iPhone 16[1].

Larangan ini diberlakukan karena *Apple* belum memenuhi persyaratan Tingkat Komponen Dalam Negeri (TKDN) sebesar 40%, yang merupakan bagian dari kebijakan pemerintah untuk mendorong investasi lokal dan meningkatkan partisipasi dalam ekonomi Indonesia. Selain itu, *Apple* juga belum merealisasikan pembangunan *Apple Developer Academy* di Bali, yang menjadi salah satu syarat utama untuk memenuhi TKDN. Kondisi ini memicu tindakan tegas dari pemerintah, termasuk rencana menonaktifkan IMEI pada perangkat yang tidak memenuhi ketentuan TKDN, sehingga perangkat tersebut tidak dapat mengakses jaringan seluler di Indonesia. Kebijakan ini juga dimaksudkan untuk mencegah

masuknya sekitar 9.000 unit iPhone 16 yang telah masuk ke Indonesia melalui jalur penumpang secara ilegal, sebagaimana tercatat oleh Kementerian Perindustrian (Kemenperin)[1].

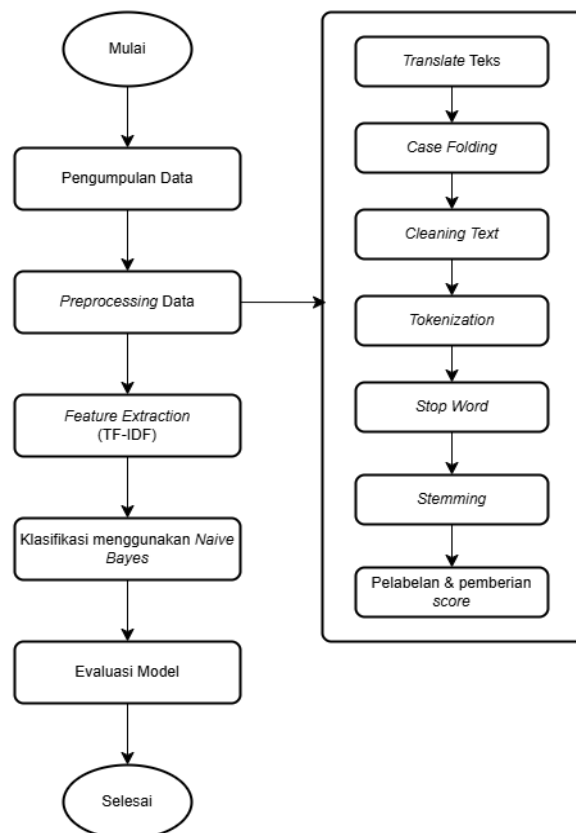
Keputusan tersebut memunculkan berbagai reaksi di media sosial, terutama di platform X. Beberapa pengguna mendukung langkah pemerintah sebagai upaya untuk melindungi industri dalam negeri sekaligus menekan peredaran produk ilegal. Namun, sebagian lainnya menganggap kebijakan ini menghambat akses masyarakat terhadap teknologi terbaru dan produk premium. Perdebatan ini semakin mencuat seiring tingginya aktivitas pengguna media sosial di Indonesia, yang membuat opini mereka memiliki potensi besar memengaruhi persepsi publik secara luas.

Untuk menganalisis sentimen masyarakat terhadap kebijakan ini, *Algoritma Naive Bayes* menjadi salah satu pendekatan yang efektif. Algoritma ini menganalisis data pelatihan dengan menghitung frekuensi setiap kata, sehingga dapat memberikan prediksi terhadap data baru. Metode ini sangat cocok untuk klasifikasi teks, terutama dalam menangani dataset berdimensi tinggi seperti komentar di media sosial. Selain itu, *Naive Bayes* dikenal memiliki tingkat kesalahan yang lebih rendah dan performa tinggi, terutama ketika diterapkan pada dataset besar, sehingga memungkinkan analisis yang cepat dan akurat dalam memahami respons publik terhadap kebijakan tersebut[2].

Penelitian ini berfokus pada analisis sentimen pengguna X terhadap kebijakan pemerintah yang melarang penjualan iPhone 16. Dengan menggunakan metode klasifikasi Naive Bayes, penelitian ini bertujuan untuk mengidentifikasi sentimen positif, negatif, dan netral yang tercermin dari berbagai tweet yang membahas topik ini. Penelitian ini dapat memberikan pemahaman yang komprehensif mengenai persepsi publik terhadap kebijakan pemerintah dalam konteks ekonomi dan industri.

II. METODE PENELITIAN

Pada penelitian ini menggunakan pendekatan kualitatif untuk menganalisis sentimen berbasis emosi terhadap kebijakan pemerintah Indonesia tentang larangan penjualan iPhone 16. Penelitian ini bertujuan untuk mengidentifikasi dan menganalisis emosi yang terkandung dalam opini publik yang diekspresikan di media sosial dan platform berita. Analisis dilakukan dengan menggunakan algoritma *Naive Bayes* untuk mengklasifikasikan sentimen menjadi kategori positif, negatif, dan netral. Berikut tahapan dari diagram alir penelitian :



Gambar 1. Diagram Alir Penelitian

A. Pengumpulan Data

Tahap awal dalam penelitian ini berfokus pada proses pengumpulan data melalui proses *crawling*. Data penelitian diperoleh dengan memanfaatkan kata kunci yang sesuai dengan topik penelitian yang bersumber dari media sosial X. Data yang dikumpulkan berupa tweet yang dipublikasikan oleh pengguna X[3].

B. *Preprocessing* Data

Preprocessing data adalah tahap pengolahan yang memiliki tujuan untuk memodifikasi data yang tidak sesuai menjadi lebih terstruktur[4]. Proses ini melibatkan penghapusan informasi yang tidak relevan atau penyesuaian data sehingga lebih mudah diolah, dan menjadi sangat krusial dalam analisis sentimen, terutama di media sosial, dimana teks seringkali bersifat informal, tidak terorganisir, dan mengandung banyak elemen yang mengganggu (*noise*)[5]. *Preprocessing* data memiliki tahapan sebagai berikut :

1. *Translate* Teks

Translate atau terjemahan adalah proses memindahkan makna suatu teks dari bahasa asal ke bahasa tujuan. Contohnya, seorang penerjemah mengubah teks dari bahasa Inggris ke bahasa Indonesia, yang berarti penerjemah mentransfer pesan atau isi dari bahasa sumber ke bahasa sasaran. Dalam konteks ini, bahasa Inggris berperan sebagai bahasa sumber, sementara bahasa Indonesia sebagai bahasa sasaran[6].

2. *Case Folding*

Case folding adalah proses yang mengubah semua huruf dalam teks menjadi huruf kecil. Tujuannya untuk menghindari perbedaan dalam pengenalan antara huruf kapital dan huruf kecil dalam data[7].

3. *Cleaning Text*

Proses *cleaning text* bertujuan untuk membersihkan data dengan menghapus elemen-elemen yang kurang relevan untuk penelitian ini. Tahapan yang dilakukan meliputi penghapusan tag, hastag, URL, mention (username), tanda baca, angka, emoji, spasi berlebih, serta pengurangan karakter yang berulang[7].

4. *Tokenization*

Tokenization adalah proses yang membagi kalimat menjadi kata-kata atau unit yang lebih kecil, sehingga memudahkan pemrosesan terkait *stopword*[8].

5. *Stopword*

Stopword adalah proses penghapusan kata-kata yang umum digunakan tetapi tidak memiliki makna signifikan. Contoh *stopword* dalam bahasa Indonesia termasuk “yang”, “dan”, “di”, “dari”, dan lainnya. Tujuannya untuk menghapus kata-kata yang kurang relevan, sehingga analisis dapat lebih terfokus pada kata-kata yang lebih penting[9].

6. *Stemming*

Stemming adalah proses mengubah kata-kata berimbuhan menjadi bentuk dasarnya berdasarkan kamus atau pustaka tertentu yang digunakan[3].

7. Pelabelan dan Pemberian *Score*

Pelabelan merupakan tahapan memberikan label pada data, tahapan ini menggunakan tiga label yaitu positif, negatif dan netral. Pemberian label pada tahapan ini dengan pemberian *score*. Jika *score* lebih dari sama dengan satu, maka labelnya positif. jika nilai *score* kurang dari 0 maka labelnya negatif. Label netral akan muncul jika nilai total adalah 0[10].

C. *Feature Extraction* (TF-IDF)

Setelah tahap *preprocessing* data, tahap selanjutnya yaitu *feature extraction* dengan menggunakan metode *Term Frequency-Inverse Document Frequency* atau biasa disebut TF-IDF. TF-IDF merupakan teknik yang sering digunakan dalam pemrosesan bahasa alami untuk memberikan bobot pada kata-kata dalam sebuah dokumen berdasarkan frekuensi kemunculannya dan seberapa unik kata tersebut dalam keseluruhan dokumen. Metode ini memiliki dua faktor utama[11]:

1. *Term Frequency* (TF): Menghitung frekuensi kemunculan sebuah kata dalam sebuah dokumen.
2. *Inverse Document Frequency* (IDF): Menilai tingkat kepentingan sebuah kata berdasarkan seberapa jarang kata tersebut muncul di seluruh dokumen.

Rumus *Term Frequency-Inverse Document Frequency* (TF-IDF):

$$tf = 0,5 + 0,5 \times \frac{tf}{\max(tf)} \quad (1)$$

$$idf_t = \log\left(\frac{D}{df_t}\right) \quad (2)$$

$$Wd.t = tf_{d,t} \times idf_{d,t} \quad (3)$$

Keterangan :

tf	= Frekuensi kemunculan suatu kata dalam sebuah dokumen
max(tf)	= Frekuensi kata yang paling sering muncul dalam dokumen
idf	= Inverse document frequency
t	= term ke t dari dokumen
D	= Jumlah total dokumen dalam kumpulan data
df _t	= Jumlah dokumen yang mengandung term t
Wd.t	= Bobot TF-IDF untuk term t di dokumen d
tf _{d,t}	= Nilai term frequency untuk term t di dokumen d
idf _{d,t}	= Nilai inverse document frequency untuk term t

D. Klasifikasi menggunakan *Naive Bayes*

Naive Bayes merupakan salah satu algoritma yang digunakan dalam teknik *data mining*. *Naive Bayes* digunakan untuk memprediksi kelas atau memberikan label dari data yang didapat berdasarkan peluang dari karakteristik yang dimilikinya. Dalam penelitian ini, diterapkan metode klasifikasi pada tahap implementasi menggunakan algoritma *Naive Bayes* untuk menguji data berdasarkan model yang dibangun dari data latih, yang digunakan untuk melatih algoritma tersebut. *Naive Bayes* memiliki rumus secara umum sebagai berikut[12] :

$$P(A|B) = \frac{P(A|B) \times P(A)}{P(B)} \quad (4)$$

Keterangan :

A	= Hipotesis data pada kelas yang spesifik.
B	= Data kelas yang belum diketahui.
P(A B)	= Probabilitas hipotesis A berdasarkan pada kondisi B.
P(A)	= Probabilitas A.
P(B A)	= Probabilitas B berdasarkan pada hipotesis H
P(B)	= Probabilitas B.

E. Evaluasi Model

Dalam mengevaluasi hasil dari algoritma *Naive Bayes*, terdapat beberapa metrik yang umum digunakan[13]:

1. Akurasi (*Accuracy*)

Tahapan ini mengukur persentase kebenaran prediksi model terhadap data yang diamati. Perbandingan antara total prediksi yang akurat dan jumlah keseluruhan prediksi dinyatakan sebagai akurasi. Berikut merupakan rumus dari akurasi:

$$accuracy = \frac{TP+TNN}{TP+FP+TN+FN} \times 100\% \quad (5)$$

2. Presisi (*Precision*)

Tahapan ini mengevaluasi tingkat kebenaran dari prediksi positif. Perhitungan presisi melibatkan pembagian antara jumlah prediksi positif yang akurat dibandingkan dengan total prediksi positif. Berikut merupakan rumus dari presisi:

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

Tahapan ini menilai sejauh mana model mampu menemukan atau mengenali kasus positif. Recall dinyatakan sebagai perbandingan antara jumlah prediksi positif yang akurat dibandingkan dengan total kasus positif yang sesungguhnya.

3. F1-Score

Menggabungkan presisi dan recall menjadi satu skor yang seimbang. F1-Score dinyatakan sebagai rata-rata harmonik antara presisi dan recall. Berikut merupakan rumus dari F1-Score:

$$f1 - score = 2 * \frac{Precision * recall}{precision + recall} \quad (7)$$

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pengumpulan data merupakan tahap awal dalam penelitian ini. Data dikumpulkan menggunakan teknik *crawling* dengan kata kunci yang relevan dengan topik penelitian dari media sosial X. Data yang diperoleh berupa *tweet* dari pengguna media sosial X dengan kata kunci berikut:

1. Iphone 16 Indonesia
2. Kebijakan pemerintah indonesia Iphone 16
3. Indonesia melarang Iphone 16

Total data yang diperoleh sebanyak 546 *tweet* yang dikumpulkan dari bulan Juli hingga November 2024. Data yang dikumpulkan terdiri dari teks berbahasa Indonesia dan Inggris. Proses pengumpulan data dilakukan menggunakan *Google Colab*, seperti ditunjukkan pada gambar 2.

	username	full_text
0	Strategi_Bisnis	Iphone 16 dilarang dijual di RI. Sebabnya blm ...
1	CNNIDdaily	Komisi 6 DPR RI mendesak BUMN sebagai perwakil...
2	asumsico	Kementerian Perindustrian dikabarkan tengah me...
3	TirtoID	Anggota Komisi VI DPR RI dari Fraksi PDIP Muft...
4	GlobeEyeNews	Indonesia bans the iPhone 16 due to Apple's un...
...
541	Gardapublik_id	Daftar Harga iPhone 15 di Indonesia *Berikut a...
542	galleryphonesby	PRODUCT SECOND/BEKAS IPHONE 14 PRO 128 SPACE B...
543	Sang_GarudaBola	Ketum PSSI @erickthohir mngatakan agenda persi...
544	galluckyx	@tanyarfes setauku mgkn krm lifetime/daya pak...
545	MacNCheeseID	@CS13231 @AbhishekMarkets In Indonesia it woul...

546 rows x 2 columns

Gambar 2. Data yang diperoleh

B. Preprocessing Data

Tahap *preprocessing* dilakukan untuk mengolah data yang telah dikumpulkan agar dapat digunakan dan memiliki nilai dalam klasifikasi. Tahapan pra-pemrosesan yang digunakan dalam penelitian ini meliputi:

1. Translate Teks

Proses ini menerjemahkan teks yang berbahasa Inggris ke dalam bahasa Indonesia. Proses penerjemahan dilakukan menggunakan pustaka `googletrans==3.1.0a0` pada platform *Google Colab*. Hasil perbandingan antara teks yang telah diterjemahkan dengan teks aslinya ditunjukkan pada gambar 3, contohnya pada data ke 4 dan 545.

	full_text	Translate_text
0	Iphone 16 dilarang dijual di RI. Sebabnya blm ...	Iphone 16 dilarang dijual di RI. Sebabnya blm ...
1	Komisi 6 DPR RI mendesak BUMN sebagai perwakil...	Komisi 6 DPR RI mendesak BUMN sebagai perwakil...
2	Kementerian Perindustrian dikabarkan tengah me...	Kementerian Perindustrian dikabarkan tengah me...
3	Anggota Komisi VI DPR RI dari Fraksi PDIP Muft...	Anggota Komisi VI DPR RI dari Fraksi PDIP Muft...
4	Indonesia bans the iPhone 16 due to Apple's un...	Indonesia melarang iPhone 16 karena komitmen i...
...
541	Daftar Harga iPhone 15 di Indonesia *Berikut a...	Daftar Harga iPhone 15 di Indonesia *Berikut a...
542	PRODUCT SECOND/BEKAS IPHONE 14 PRO 128 SPACE B...	PRODUK SECOND/BEKAS IPHONE 14 PRO 128 SPACE BL...
543	Ketum PSSI @erickthohir mngatakan agenda persi...	Ketum PSSI @erickthohir mngatakan agenda persi...
544	@lanyarfes setauku mgkn krn lifetime/daya pak...	@lanyarfes setauku mgkn krn seumur hidup/daya...
545	@CS13231 @AbhishekMarkets In Indonesia it woul...	@CS13231 @AbhishekMarkets Di Indonesia harganya...

546 rows x 2 columns

Gambar 3. Hasil Terjemahan Teks

2. Case Folding

Proses ini mengubah semua huruf dalam teks menjadi huruf kecil (*case folding*). Proses *case folding* dilakukan menggunakan *Google Colab*. Hasil perbandingan antara teks yang telah diubah menjadi huruf kecil dengan hasil translate ditunjukkan pada gambar 4.

	Translate_text	Casefolding_text
0	Iphone 16 dilarang dijual di RI. Sebabnya blm ...	iphone 16 dilarang dijual di ri. sebabnya blm ...
1	Komisi 6 DPR RI mendesak BUMN sebagai perwakil...	komisi 6 dpr ri mendesak bumn sebagai perwakil...
2	Kementerian Perindustrian dikabarkan tengah me...	kementerian perindustrian dikabarkan tengah me...
3	Anggota Komisi VI DPR RI dari Fraksi PDIP Muft...	anggota komisi vi dpr ri dari fraksi pdip muft...
4	Indonesia melarang iPhone 16 karena komitmen i...	indonesia melarang iphone 16 karena komitmen i...
...
541	Daftar Harga iPhone 15 di Indonesia *Berikut a...	daftar harga iphone 15 di indonesia *berikut a...
542	PRODUK SECOND/BEKAS IPHONE 14 PRO 128 SPACE BL...	produk second/bekas iphone 14 pro 128 space bl...
543	Ketum PSSI @erickthohir mngatakan agenda persi...	ketum pssi @erickthohir mngatakan agenda persi...
544	@lanyarfes setauku mgkn krn seumur hidup/daya...	@lanyarfes setauku mgkn krn seumur hidup/daya...
545	@CS13231 @AbhishekMarkets Di Indonesia harganya...	@cs13231 @abhishekmarkets di indonesia harganya...

546 rows x 2 columns

Gambar 4. Hasil Perbandingan Case Folding dengan Translate Teks

3. Cleaning Text

Proses ini membersihkan teks dengan menghapus elemen-elemen seperti *mention*, *hashtag*, URL, dan simbol-simbol yang tidak relevan. Proses *cleaning text* dilakukan menggunakan platform *Google Colab*. Hasil perbandingan antara teks setelah proses *cleaning text* dengan teks setelah *case folding* ditunjukkan pada gambar 5.

	Casefolding_text	cleaning_text
0	iphone 16 dilarang dijual di ri. sebabnya blm ...	iphone dilarang dijual di ri sebabnya blm mem...
1	komisi 6 dpr ri mendesak bumh sebagai perwakil...	komisi dpr ri mendesak bumh sebagai perwakila...
2	kementerian perindustrian dikabarkan tengah me...	kementerian perindustrian dikabarkan tengah me...
3	anggota komisi vi dpr ri dari fraksi pdip muft...	anggota komisi vi dpr ri dari fraksi pdip muft...
4	indonesia melarang iphone 16 karena komitmen i...	indonesia melarang iphone karena komitmen inv...
...
541	daftar harga iphone 15 di indonesia *berikut a...	daftar harga iphone di indonesia berikut adal...
542	produk second/bekas iphone 14 pro 128 space bl...	produk secondbekas iphone pro space black ib...
543	ketum pssi @erickthohir mngatakan agenda persi...	ketum pssi mngatakan agenda persiapan timnas ...
544	@tanyarifes setauku mgkn krm seumur hidup/daya...	setauku mgkn krm seumur hidupdaya pakai iphone...
545	@cs13231 @abhishekmarkets di indonesia hargany...	di indonesia harganya rp usd vs iphone ber...

546 rows x 2 columns

Gambar 5. Hasil Perbandingan *Cleaning Text* dengan *Case Folding*

4. Tokenization

Proses ini memecah atau membagi kalimat menjadi kata-kata atau token-token (*tokenization*). Proses *tokenization* dilakukan menggunakan platform *Google Colab* dan pustaka NLTK. Hasil perbandingan antara teks hasil *tokenization* dengan teks setelah proses *cleaning* ditunjukkan pada gambar 6.

	cleaning_text	tokenized
0	iphone dilarang dijual di ri sebabnya blm mem...	[iphone, dilarang, dijual, di, ri, sebabnya, b...
1	komisi dpr ri mendesak bumh sebagai perwakila...	[komisi, dpr, ri, mendesak, bumh, sebagai, per...
2	kementerian perindustrian dikabarkan tengah me...	[kementerian, perindustrian, dikabarkan, tanga...
3	anggota komisi vi dpr ri dari fraksi pdip muft...	[anggota, komisi, vi, dpr, ri, dari, fraksi, p...
4	indonesia melarang iphone karena komitmen inv...	[indonesia, melarang, iphone, karena, komitmen...
...
541	daftar harga iphone di indonesia berikut adal...	[daftar, harga, iphone, di, indonesia, berikut...
542	produk secondbekas iphone pro space black ib...	[produk, secondbekas, iphone, pro, space, blac...
543	ketum pssi mngatakan agenda persiapan timnas ...	[ketum, pssi, mngatakan, agenda, persiapan, ti...
544	setauku mgkn krm seumur hidupdaya pakai iphone...	[setauku, mgkn, krm, seumur, hidupdaya, pakai,...
545	di indonesia harganya rp usd vs iphone ber...	[di, indonesia, harganya, rp, usd, vs, iphone,...

546 rows x 2 columns

Gambar 6. Hasil Perbandingan Teks *Tokenized* dengan *Cleaning Text*

5. Stopword

Proses ini menghapus kata-kata yang tidak penting (*stopwords*) yang tidak memberikan kontribusi signifikan terhadap analisis, seperti "di", "yang", atau "dan". Proses penghapusan *stopwords* dilakukan menggunakan platform *Google Colab* dan pustaka Sastrawi. Hasil perbandingan antara teks setelah penghapusan *stopwords* dengan teks setelah proses *tokenization* ditunjukkan pada gambar 7.

	tokenized	stopword
0	[iphone, dilarang, dijual, di, ri, sebabnya, b...	[iphone, dilarang, dijual, , ri, sebabnya, blm...
1	[komisi, dpr, ri, mendesak, bumh, sebagai, per...	[komisi, dpr, ri, mendesak, bumh, , perwakilan...
2	[kementerian, perindustrian, dikabarkan, tenga...	[kementerian, perindustrian, dikabarkan, tenga...
3	[anggota, komisi, vi, dpr, ri, dari, fraksi, p...	[anggota, komisi, vi, dpr, ri, , fraksi, pdip,...
4	[indonesia, melarang, iphone, karena, komitmen...	[indonesia, melarang, iphone, , komitmen, inve...
...
541	[daftar, harga, iphone, di, indonesia, berikut...	[daftar, harga, iphone, , indonesia, berikut, ...
542	[produk, secondbekas, iphone, pro, space, blac...	[produk, secondbekas, iphone, pro, space, blac...
543	[ketum, pssi, mngatakan, agenda, persiapan, ti...	[ketum, pssi, mngatakan, agenda, persiapan, ti...
544	[setauku, mgkn, km, seumur, hidupdaya, pakai,...	[setauku, mgkn, km, seumur, hidupdaya, pakai,...
545	[di, indonesia, harganya, rp, usd, vs, iphone,...	[, indonesia, harganya, rp, usd, vs, iphone, b...

546 rows x 2 columns

Gambar 7. Hasil Perbandingan Teks *Stopword* dengan *Tokenized*

6. Stemming

Proses ini mengubah kata-kata ke bentuk dasarnya (*stemming*). Proses *stemming* dilakukan menggunakan platform *Google Colab* dan pustaka *Sastrawi*, yang memiliki kamus kata dasar bahasa Indonesia. Hasil perbandingan antara teks setelah proses *stemming* dengan teks setelah penghapusan *stop words* ditunjukkan pada gambar 8.

	stopword	stemming
0	[iphone, dilarang, dijual, , ri, sebabnya, blm...	[iphone, larang, jual, , ri, sebab, blm, penuh...
1	[komisi, dpr, ri, mendesak, bumh, , perwakilan...	[komisi, dpr, ri, desak, bumh, , wakil, perint...
2	[kementerian, perindustrian, dikabarkan, tenga...	[menteri, industri, kabar, tengah, serta, edar...
3	[anggota, komisi, vi, dpr, ri, , fraksi, pdip,...	[anggota, komisi, vi, dpr, ri, , fraksi, pdip,...
4	[indonesia, melarang, iphone, , komitmen, inve...	[indonesia, larang, iphone, , komitmen, invest...
...
541	[daftar, harga, iphone, , indonesia, berikut, ...	[daftar, harga, iphone, , indonesia, ikut, , d...
542	[produk, secondbekas, iphone, pro, space, blac...	[produk, secondbekas, iphone, pro, space, blac...
543	[ketum, pssi, mngatakan, agenda, persiapan, ti...	[tum, pssi, mngatakan, agenda, siap, timnas, i...
544	[setauku, mgkn, km, seumur, hidupdaya, pakai,...	[tau, mgkn, km, umur, hidupdaya, pakai, iphon...
545	[, indonesia, harganya, rp, usd, vs, iphone, b...	[, indonesia, harga, rp, usd, vs, iphone, harg...

546 rows x 2 columns

Gambar 8. Hasil Perbandingan Teks *Stemming* dengan *Stopword*

Setelah tahap *preprocessing*, dilakukan penggabungan dari setiap token. Hasil teks setelah proses *preprocessing* ditunjukkan pada gambar 9.

	full_text	processed_text
0	Iphone 16 dilarang dijual di RI. Sebabnya blm ...	iphone larang jual ri sebab blm penuh syarat ...
1	Komisi 6 DPR RI mendesak BUMN sebagai perwakil...	komisi dpr ri desak bumn wakil perintah tola...
2	Kementerian Perindustrian dikabarkan tengah me...	menteri industri kabar tengah serta edar iphon...
3	Anggota Komisi VI DPR RI dari Fraksi PDIP Muft...	anggota komisi vi dpr ri fraksi pdip mufti ai...
4	Indonesia bans the iPhone 16 due to Apple's un...	indonesia larang iphone komitmen investasi ap...
...
541	Daftar Harga iPhone 15 di Indonesia *Berikut a...	daftar harga iphone indonesia ikut daftar ha...
542	PRODUCT SECOND/BEKAS IPHONE 14 PRO 128 SPACE B...	produk secondbekas iphone pro space black ibox...
543	Ketum PSSI @erickthohir mngatakan agenda persi...	tum pssi mngatakan agenda siap timnas indonesi...
544	@tanyarifes setaoku mgkn km lifetime/daya pak...	tau mgkn km umur hidupdaya pakai iphone tuh ...
545	@CS13231 @AbhishekMarkets In Indonesia it woul...	indonesia harga rp usd vs iphone harga rp usd

546 rows x 2 columns

Gambar 9. Hasil *Preprocessing Text*

7. Pelebelan dan Pemberian *Score*

Selanjutnya, dilakukan pelabelan pada data teks menggunakan *SenticNet* dan *lexicon* untuk mengklasifikasikan atau memberi kategori pada setiap teks yang telah diproses. Dengan menggunakan *SenticNet*, proses pelabelan dapat mencakup penentuan sentimen positif, negatif, atau netral berdasarkan kata-kata yang terdapat dalam teks. Selain itu, *lexicon* dapat digunakan untuk memperkaya proses pelabelan dengan memanfaatkan kamus kata-kata yang mengandung informasi mengenai makna dan asosiasi kata-kata tersebut dalam konteks yang lebih luas. Proses pelabelan ini bertujuan untuk memberikan label yang relevan pada setiap teks. Hasil dari pelabelan dan pemberian skor ditunjukkan pada gambar 10, dan jumlah label yang didapat ditunjukkan pada gambar 11.

	full_text	processed_text	polarity	polarity_score
0	Iphone 16 dilarang dijual di RI. Sebabnya blm ...	iphone larang jual ri sebab blm penuh syarat ...	negative	-1.949
1	Komisi 6 DPR RI mendesak BUMN sebagai perwakil...	komisi dpr ri desak bumn wakil perintah tola...	neutral	0.595
2	Kementerian Perindustrian dikabarkan tengah me...	menteri industri kabar tengah serta edar iphon...	positive	2.371
3	Anggota Komisi VI DPR RI dari Fraksi PDIP Muft...	anggota komisi vi dpr ri fraksi pdip mufti ai...	positive	2.811
4	Indonesia bans the iPhone 16 due to Apple's un...	indonesia larang iphone komitmen investasi ap...	negative	-0.811
...
541	Daftar Harga iPhone 15 di Indonesia *Berikut a...	daftar harga iphone indonesia ikut daftar ha...	positive	3.291
542	PRODUCT SECOND/BEKAS IPHONE 14 PRO 128 SPACE B...	produk secondbekas iphone pro space black ibox...	positive	5.642
543	Ketum PSSI @erickthohir mngatakan agenda persi...	tum pssi mngatakan agenda siap timnas indonesi...	neutral	0.885
544	@tanyarifes setaoku mgkn km lifetime/daya pak...	tau mgkn km umur hidupdaya pakai iphone tuh ...	positive	2.547
545	@CS13231 @AbhishekMarkets In Indonesia it woul...	indonesia harga rp usd vs iphone harga rp usd	negative	0.000

546 rows x 4 columns

Gambar 10. Hasil Pelabelan dan Pemberian *Score*

	Sentimen	Jumlah	Persentase (%)
0	Positif	217	39.74
1	Negatif	181	33.15
2	Netral	148	27.11

Gambar 11. Jumlah Label yang didapat

C. Feature Extraction

Setelah tahap pelabelan, dilakukan *Feature Extraction* untuk mengubah teks yang telah diproses menjadi representasi numerik yang dapat digunakan oleh model klasifikasi. Metode yang digunakan dalam *Feature Extraction* adalah TF-IDF (*Term Frequency-Inverse Document Frequency*) yang membantu dalam mengidentifikasi kata-kata yang memiliki makna penting dalam konteks dokumen, serta membedakan kata-kata yang lebih relevan dalam analisis. Hasil dari *Feature Extraction* dengan menggunakan TF-IDF akan digunakan sebagai input untuk model klasifikasi yang akan diterapkan pada data teks. Hasil dari *Feature Extraction* menggunakan TF-IDF ditunjukkan pada gambar 12.

	aapl	aar	ada	adain	adapter	adek	agakagak	agenda	agus	agustusoktober	...	youtz	yuk	yung	yunho	yups	yuta	zenfone	zonk	zoomnya	zpa		
0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
1	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	0.0	0.0	0.226148	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
3	0.0	0.0	0.205396	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
...
541	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
542	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
543	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.289761	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
544	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
545	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	

546 rows x 1414 columns

Gambar 12. Hasil TF-IDF

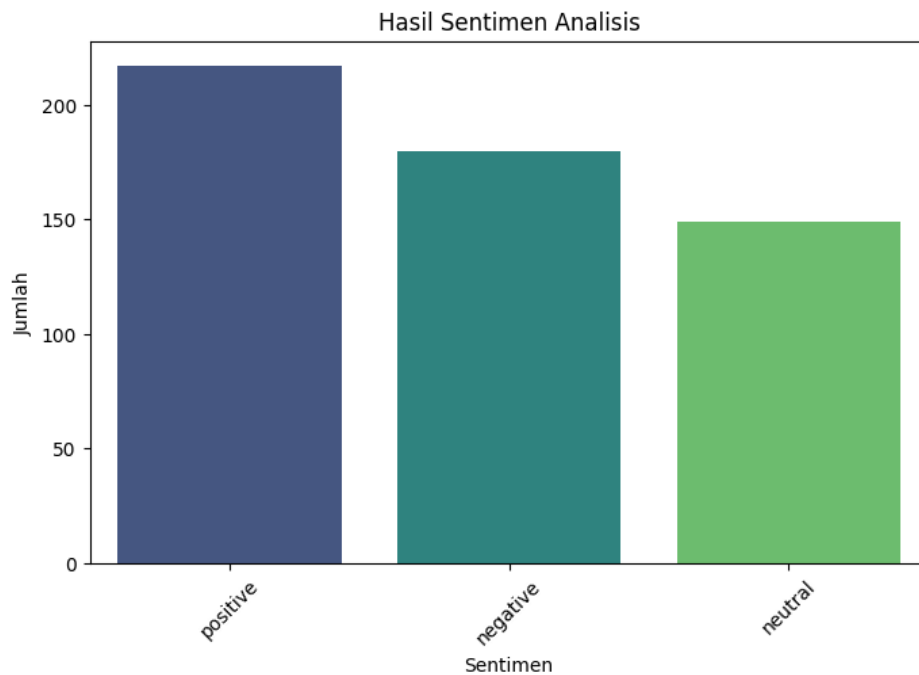
D. Klasifikasi menggunakan *Naive Bayes*

Setelah proses ekstraksi fitur menggunakan TF-IDF, langkah selanjutnya adalah melakukan klasifikasi dengan *Naive Bayes* untuk mengidentifikasi sentimen dalam teks yang telah diproses. Jika ketidakseimbangan data tidak mempengaruhi akurasi prediksi, maka model masih bisa memberikan hasil yang baik tanpa memerlukan oversampling. *Naive Bayes* akan mengklasifikasikan setiap teks ke dalam kategori sentimen, yang terdiri dari positif, negatif, atau netral. Hasil akurasi dan perhitungan *naive bayes* ditunjukkan pada gambar 13.

Akurasi: 0.73				
Hasil Klasifikasi:				
	precision	recall	f1-score	support
negative	0.78	0.72	0.75	39
neutral	0.87	0.57	0.69	35
positive	0.63	0.89	0.74	36
accuracy			0.73	110
macro avg	0.76	0.73	0.72	110
weighted avg	0.76	0.73	0.72	110

Gambar 13. Hasil akurasi dan perhitungan *naive bayes*

Untuk memvisualisasikan hasil sentimen yang diperoleh dari klasifikasi *Naive Bayes*, dibuat grafik batang yang menggambarkan distribusi sentimen dalam dataset. Grafik ini menunjukkan jumlah teks yang termasuk dalam masing-masing kategori sentimen (positif, negatif, netral). Berikut adalah gambar yang menunjukkan distribusi sentimen dalam dataset, yang dapat dilihat pada Gambar 14.



Gambar 14. Hasil Klasifikasi Sentimen

Hasil klasifikasi menggunakan *naive bayes* lebih diunggulkan dikarenakan lebih fleksibel dan sangat cocok digunakan pada dataset dalam bentuk teks. Sedangkan pemberian *score* memerlukan konversi dari teks menjadi *score* yang sulit jika data berupa kalimat panjang dan cocok jika data sudah memiliki atribut *score* seperti rating produk.

E. Evaluasi Model

Setelah melatih model *Naive Bayes*, langkah selanjutnya adalah melakukan evaluasi kinerja model menggunakan beberapa metrik evaluasi, yaitu Akurasi, Precision, Recall, dan F1-Score dengan menggunakan 436 data latih dan 110 data uji. Metrik-metrik ini memberikan gambaran yang lebih komprehensif mengenai sejauh mana model berhasil dalam melakukan klasifikasi dengan akurat.

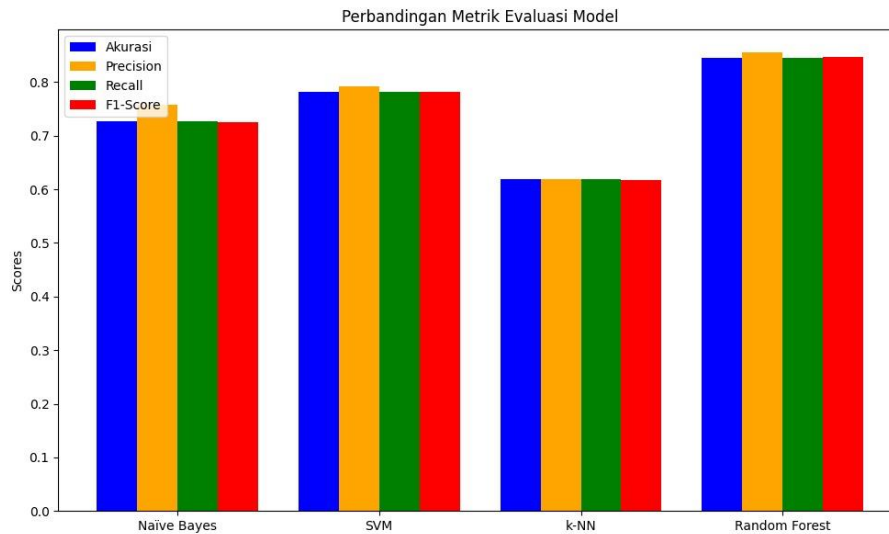
```

Hasil Evaluasi Model:
      Metric      Score
0  Accuracy  0.718182
1  Precision  0.749023
2   Recall   0.718182
3  F1-Score  0.714721
  
```

Gambar 15. Hasil Metrik Evaluasi Model

Hasil metrik evaluasi dijalankan menggunakan google colab yang ditunjukkan pada gambar 15, menunjukkan bahwa model memiliki:

1. Akurasi sebesar 0,7182, yang menunjukkan bahwa prediksi yang dihasilkan oleh model adalah benar.
2. *Precision* model sebesar 0,7490 menunjukkan bahwa dari seluruh prediksi yang diklasifikasikan sebagai positif.
3. *Recall* model sebesar 0,7182 mengindikasikan bahwa model mampu mengidentifikasi semua data yang sebenarnya positif.
4. *F1-Score* sebesar 0,7147 menunjukkan bahwa model telah mencapai keseimbangan yang baik antara akurasi dalam mengidentifikasi data positif dan menghindari kesalahan klasifikasi.



Gambar 16. Perbandingan Metrik Evaluasi Model

Secara keseluruhan, hasil evaluasi ini menunjukkan bahwa model *Naive Bayes* memiliki performa yang cukup baik dalam klasifikasi sentimen dibandingkan dengan model KNN yang ditunjukkan pada gambar 16, dengan akurasi yang memadai dan keseimbangan yang baik antara precision dan recall. Meskipun demikian, ada ruang untuk perbaikan lebih lanjut, terutama dalam meningkatkan recall dan mengurangi kesalahan prediksi, agar model dapat lebih efektif dalam menangani data yang lebih kompleks atau dataset dengan distribusi kelas yang tidak seimbang.

IV. KESIMPULAN

Penelitian ini menunjukkan bahwa masyarakat memiliki berbagai pandangan tentang larangan pemerintah terhadap penjualan iPhone 16. Sebagian mendukung kebijakan ini karena melindungi industri lokal, sementara yang lain merasa dirugikan karena akses ke teknologi terbaru menjadi terbatas. Dengan menggunakan algoritma *Naive Bayes*, sentimen masyarakat di media sosial dapat dikategorikan sebagai positif, negatif, atau netral. Akurasi model yang digunakan cukup baik, mencapai 71,82%. Namun, ada ruang untuk perbaikan agar model lebih akurat dan bisa menangani data media sosial yang rumit. Penelitian ini bermanfaat untuk memahami opini masyarakat secara luas dan bisa menjadi masukan bagi pemerintah untuk membuat kebijakan yang lebih efektif dan diterima oleh masyarakat.

UCAPAN TERIMAKASIH

Dengan penuh rasa syukur, kami mengucapkan terima kasih kepada Telkom University Purwokerto atas dukungan yang telah diberikan selama proses penelitian ini. Ucapan terima kasih juga kami sampaikan kepada semua pihak yang telah berkontribusi, baik melalui masukan maupun arahan yang sangat berharga dalam penulisan jurnal ini. Tak lupa, apresiasi kami berikan kepada keluarga dan sahabat atas motivasi dan dukungan tanpa henti. Kami berharap hasil dari penelitian ini dapat memberikan sumbangsih yang berarti bagi perkembangan ilmu pengetahuan, khususnya di bidang analisis sentimen serta penerapan algoritma *Naive Bayes*.

DAFTAR PUSTAKA

- [1] M. Taruk, A. Septiarini, F. Alief, dan A. Akbar, "Analisis Sentimen Terhadap Kebijakan Pemerintah Terkait Pandemi Covid-19 Pada Twitter," 2023.
- [2] R. Rifaldi, J. Indra, A. R. Pratama, dan A. R. Juwita, "Analisis Sentimen Pemboikotan Produk dengan Pendekatan Algoritma Naive Bayes Media Sosial X," *Journal of Information System Research (JOSH)*, vol. 5, no. 4, hlm. 1–7, 2024.
- [3] N. Q. Rizkina dan F. N. Hasan, "Analisis Sentimen Komentar Netizen Terhadap Pembubaran Konser NCT 127 Menggunakan Metode Naive Bayes," *Journal of Information System Research (JOSH)*, vol. 4, no. 4, hlm. 1–9, Jul 2023.

-
- [4] K. D. Indarwati dan H. Februariyanti, "Analisis Sentimen Terhadap Kualitas Pelayanan Aplikasi Go-jek Menggunakan Metode Naive Bayes Classifier," *JatISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, vol. 10, no. 1, hlm. 1–12, 2023.
- [5] Widia, Z. Y. Aqsalia, S. Sari, N. U. Khoirunnisa, dan F. Kurniawan, "Optimasi Algoritma Naive Bayes Untuk Menganalisis Sentimen Pada Konten Pemindahan Ibu Kota di Youtube," *Journal of Computer and Information Systems Ampera*, vol. 5, no. 2, hlm. 1–16, 2024.
- [6] J. Penelitian Teknologi Pendidikan, T. Surgawi, dan H. Joebagio, "TEKNODIKA Analisis Kualitas Terjemahan dari Bahasa Inggris ke dalam Bahasa Indonesia," 2018.
- [7] U. Khairani, V. Mutiawani, dan H. Ahmadian, "Pengaruh Tahapan Preprocessing Terhadap Model Indobert Dan Indobertweet Untuk Mendeteksi Emosi Pada Komentar Akun Berita Instagram," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 4, hlm. 1–10, Agu 2024.
- [8] T. A. Sari, E. Sinduningrum, dan F. N. Hasan, "Analisis Sentimen Ulasan Pelanggan Pada Aplikasi Fore Coffee Menggunakan Metode Naive Bayes," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 3, no. 6, hlm. 1–7, 2023.
- [9] R. Slamet, W. Gata, A. Novtariy, K. Hilyati, dan F. A. Jariyah, "Analisis Sentimen Twitter Terhadap Penggunaan Artis Korea Selatan Sebagai Brand Ambassador Produk Kecantikan Lokal," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 5, no. 1, hlm. 1–9, 2022.
- [10] K. Cindy Astuti, A. Firmansyah, A. Riyadi, dan U. Pelita Bangsa Bekasi, "Implementasi Text Mining untuk Analisis Sentimen Masyarakat terhadap Ulasan Aplikasi Digital Korlantas Polri pada Google Play Store," *Remik: Riset dan E-Jurnal Manajemen Informatika Komputer*, vol. 8, no. 1, 2024.
- [11] D. Septiani dan I. Isabela, "SINTESIA: Jurnal Sistem dan Teknologi Informasi Indonesia "Analisis Term Frequency Inverse Document Frequency (TF-IDF) dalam Temu Kembali Informasi Pada Dokumen Teks. Vol. 01, No. 2, Maret 2022.
- [12] K. Fadli dan F. N. Hasan, "Analisis Sentimen Tanggapan Masyarakat Terhadap Penutupan TikTok Shop Menggunakan Metode Naive Bayes," *Journal of Information System Research (JOSH)*, vol. 6, no. 1, hlm. 1–10, 2024.
- [13] R. Alfarez, V. Purwayoga, J. Siliwangi No, K. Kahuripan, K. Tawang, dan K. Tasikmalaya, "Penerapan Naive Bayes Untuk Prediksi Customer CHURN (Studi Kasus: PT Hutchison 3 Indonesia)," *Jurnal Riset dan Aplikasi Mahasiswa Informatika (JRAMI)*, vol. 05, 2024.