

Implementasi Metode Naïve Bayes Classifier Terhadap Analisis Sentimen Boikot Produk Terafiliasi Israel Pada Media Sosial “X”

1st Miranda Sekar Ayu Shaumi

Teknik Informatika

Science Goup

Purwokerto, Indonesia

mirandasekar@student.telkomuniversity.ac.id

2nd Paradise

Teknik Informatika

Science Goup

Purwokerto, Indonesia

paradise@telkomuniversity.ac.id

Abstrak — Konflik antara Israel dan Palestina menimbulkan dampak global, termasuk memicu gerakan pemboikotan terhadap produk-produk yang memiliki keterkaitan dengan Israel. Di Indonesia, gerakan ini mendapatkan dukungan luas, terutama di media sosial “X”. Penelitian ini dilakukan untuk mengevaluasi pandangan masyarakat terhadap gerakan tersebut melalui metode analisis sentimen berbasis teks, dengan memanfaatkan data berupa komentar pada unggahan pengguna yang kemudian dikategorikan ke dalam sentimen positif dan negatif. Isu boikot produk Israel menjadi perhatian masyarakat luas, ditandai dengan masifnya penyebaran opini di media sosial. Penelitian ini penting karena dapat memberi pemahaman tentang persepsi masyarakat yang mampu dimanfaatkan oleh pemilik merek untuk menjaga reputasi mereka. Proses analisis dijalankan dengan menerapkan algoritma Naïve Bayes Classifier terhadap data teks dari media sosial “X” sepanjang tahun 2024. Hasil klasifikasi mengindikasikan bahwasanya dari total 9.442 komentar, sebanyak 63,1% termasuk sentimen positif dan 36,9% tergolong negatif. Model mencapai tingkat akurasi sejumlah 83%, dengan nilai f1-score tertinggi pada kelas sentimen positif, yakni 0,87%. Kontribusi utama dari penelitian ini adalah penerapan pipeline klasifikasi yang efektif untuk menangkap opini publik terhadap isu sosial yang sensitif.

Kata kunci— Analisis Sentimen, Naive Bayes, Boikot, Produk Israel, Media Sosial, Klasifikasi Teks

I. PENDAHULUAN

Konflik antara Israel dan Palestina merupakan permasalahan berkepanjangan di kawasan Timur Tengah. Sejarah mencatat bahwa sebelum berdirinya negara Israel, komunitas Yahudi merupakan kelompok tanpa negara yang tersebar di berbagai wilayah, khususnya Eropa[1]. Dukungan dari pemerintah Inggris menjadi titik awal pemindahan populasi Yahudi ke wilayah Palestina, yang akhirnya menimbulkan konflik teritorial yang masih berlangsung hingga kini. Pembagian wilayah Palestina dan Israel salah satunya diatur melalui Perjanjian Oslo yang disepakati oleh Organisasi Pembebasan Palestina (PLO)[2].

Dalam beberapa tahun terakhir, ketegangan meningkat drastis, terutama setelah serangan besar-besaran oleh HAMAS pada 7 Oktober 2023 yang memicu aksi balasan oleh Israel dan menyebabkan lebih dari 33.000 korban jiwa

di Jalur Gaza[3]. Aksi militer Israel mengundang kecaman internasional, karena dianggap sebagai kejahatan perang. Sebagai bentuk solidaritas, masyarakat dunia—termasuk Indonesia—melakukan berbagai aksi, salah satunya melalui gerakan boikot terhadap produk-produk yang terafiliasi dengan Israel[3].

Majelis Ulama Indonesia (MUI) turut mengeluarkan Fatwa Nomor 83 Tahun 2023 yang mengharamkan dukungan langsung maupun tidak langsung terhadap agresi Israel, termasuk melalui konsumsi produk dari perusahaan yang mendukung tindakan tersebut[4]. Gerakan boikot ini kemudian menjadi isu yang ramai diperbincangkan di media sosial. Platform seperti Facebook, TikTok, Instagram, dan terutama “X” (sebelumnya Twitter) menjadi saluran utama penyebaran informasi, kampanye boikot, dan penggalangan opini publik[5]. Media sosial “X” menjadi pusat perhatian karena jangkauan luas dan fitur interaktifnya, dengan jumlah pengguna aktif di Indonesia mencapai 27 juta pada Oktober 2023[6]. Beberapa merek yang diketahui memiliki afiliasi atau menunjukkan dukungan terhadap Israel antara lain adalah McDonald’s dan Starbuck[7].

Melihat tingginya aktivitas opini publik di media sosial, diperlukan analisis yang dapat memahami persepsi masyarakat secara komputasional[8]. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap unggahan terkait boikot produk yang terafiliasi dengan Israel di media sosial “X”. Metode yang digunakan adalah Naïve Bayes Classifier, salah satu teknik klasifikasi berbasis probabilitas yang sering digunakan dalam analisis teks[9]. Berdasarkan penelitian terdahulu, metode ini menunjukkan akurasi tinggi dalam mengklasifikasikan sentimen positif dan negatif. Oleh karena itu, penerapan Naïve Bayes dalam konteks ini diharapkan mampu memberikan gambaran objektif mengenai sentimen publik terhadap gerakan boikot di Indonesia.

II. KAJIAN TEORI

A. Media Sosial “X”

Platform media sosial berbasis internet, pengguna media sosial “X” dapat berbagi informasi, pendapat maupun berbincang mengaplikasikan pesan singkat pada media sosial tersebut. Maksimal karakter dalam setiap unggahan yakni 280 karakter, yang dalam setiap unggahan dapat disisipkan

gambar, video maupun link. Secara global jumlah pengguna aktif media sosial X mencapai 421 juta pada akhir 2023, dan Indonesia memiliki 27 juta pengguna X pada oktober 2023. Fitur utama yang dimiliki media sosial X yakni post, repost, like, hashtags, mention, direct message, dan communities. Post merupakan pesan singkat yang mampu berisi teks, gambar, video, GIF, dan link. Repost membagikan post orang lain ke pengikut pengguna.

B. Tweet Harvest

TweetHarvest merupakan proses pengumpulan data berupa unggahan dari platform media sosial “X” secara otomatis dengan memanfaatkan Application Programming Interface (API). Proses ini bertujuan untuk memperoleh data dalam jumlah besar yang selaras dengan topik penelitian, berdasarkan kriteria tertentu seperti kata kunci dan rentang waktu. Data hasil tweet harvest umumnya digunakan pada keperluan analisis lebih lanjut, seperti analisis sentimen, identifikasi tren, atau kajian perilaku pengguna di media sosial.

C. Machine Learning

Machine learning ialah teknologi kecerdasan buatan dimana mesin memungkinkan untuk mempelajari sesuatu melalui data tanpa adanya pemrograman secara eksplisit. Algoritma machine learning memberi kemungkinan bagi sistem untuk mengenali pola dan kecenderungan dalam data guna menghasilkan keputusan yang tepat dan berbasis data[10]. Algoritma *Machine Learning* merupakan serangkaian aturan atau instruksi yang mengarahkan proses pembelajaran. Terdapat tiga jenis utama dalam *machine learning*, yakni *supervised learning*, *unsupervised learning*, dan *reinforcement learning*[11].

D. Analisis Sentimen

Analisis sentimen merupakan teknik untuk menganalisis sentiment atau perasaan yang terkandung dalam teks[12]. Melalui analisis sentimen, informasi berharga dapat didapat dari data yang tidak terstruktur, di mana opini yang terkandung di dalamnya dapat diklasifikasikan ke dalam sentimen positif atau negatif. Klasifikasi ini didasari muatan emosional yang terkandung dalam opini tersebut[13]. Pengelompokan sentimen tersebut dapat dilakukan mengaplikasikan klasifikasi teks[14].

Tahapan analisis sentimen yakni:

- Crawling Data
- Preprocessing Data (Case Folding, Cleaning, Tokenisasi, Stemming, Stopward Removal)
- Pelabelan
- Visualisasi Data
- Tf-IDF dan Kasifikasi Model
- Evaluasi Confusion Matrix

E. Algoritma Naïve Bayes Classifier

Naïve Bayes Classifier merupakan metode klasifikasi statistik yang diaplikasikan untuk memperkirakan probabilitas suatu data tergolong ke dalam kelas tertentu. Algoritma ini berlandaskan pada Teorema Bayes dan memiliki kemampuan klasifikasi yang setara dengan metode lain, seperti decision tree maupun neural network. Selain itu, Naïve Bayes Classifier dikenal memiliki tingkat akurasi yang tinggi serta efisiensi dalam proses komputasi, terutama ketika diterapkan pada basis data berukuran besar[15].

$$P(A|B) = \frac{P(B \cap A)}{P(B)} = \frac{P(B|A) \cdot P(A)}{\sum_{i=1}^n P(B|A_i) \cdot P(A_i)} \quad (1)$$

Keterangan:

A = hipotesis data B merupakan suatu class spesifik

B = data dengan class yang belum diketahui

P(A) = Probabilitas awal dari A sebelum ada bukti B (disebut juga prior probability).

P(B) = Probabilitas bahwa kejadian B terjadi (disebut juga evidence).

P(A|B) = Probabilitas hipotesis A benar jika diketahui bahwa kejadian B telah terjadi (disebut juga posterior probability).

P(B|A) = Probabilitas bahwa B terjadi jika A benar (disebut juga likelihood).

P(B∩A) = Probabilitas bahwa A dan B terjadi secara bersamaan (irisan A dan B).

F. Confusion Matrix

Confusion Matrix adalah metode yang umum diterapkan guna menghitung metrik evaluasi seperti akurasi, recall, precision, dan tingkat kesalahan (error). Dalam evaluasi recall, dilakukan proses identifikasi terhadap aplikasi guna menemukan seluruh item yang selaras dengan dokumen. Confusion Matrix berfungsi untuk mengilustrasikan hasil klasifikasi model terhadap data uji, yang dikelompokkan ke dalam empat kategori, yakni True Positive (TP), False Positive (FP), False Negative (FN), dan True Negative (TN)[9].

H. Phyton

Python merupakan bahasa pemrograman serbaguna yang umumnya diterapkan guna pengembangan situs, perangkat lunak, otomatisasi tugas, dan analisis data. Sebagai bahasa pemrograman tujuan umum, python dapat diterapkan guna menciptakan berbagai jenis program tanpa terbatas pada masalah tertentu.

I. Google Colab

Google Collaboratory, yang lebih dikenal dengan sebutan Google Colab, adalah sebuah platform berbasis cloud yang memungkinkan pengguna menulis, menjalankan, dan membagikan kode Python secara langsung melalui browser web. Platform ini dirancang khusus untuk analis, pengembang, peneliti, serta pendidik di bidang data science dan machine learning, dengan menyediakan fleksibilitas pada lingkungan komputasi, kemudahan akses, serta dapat digunakan secara gratis[15].

III. METODE

Dalam penelitian ini diperlukan beberapa tahapan yang diperlukan, berikut diagram alir penelitian tertera pada Gambar 1.



GAMBAR 1
DIAGRAM ALIR PENELITIAN

A. Studi Literatur

Pada tahap ini, peneliti melakukan pencarian literatur yang relevan dengan topik penelitian. Proses pencarian dilakukan melalui sumber-sumber ilmiah terpercaya, dengan memanfaatkan Google Scholar sebagai media utama, serta menggunakan aplikasi Publish or Perish untuk menjangkau artikel dalam rentang lima tahun terakhir.

B. Perumusan Masalah

Selanjutnya di tahap ini, masalah terumuskan dengan mengamati sejumlah topik yang sedang tren di media sosial X, khususnya yang berkaitan dengan ajakan untuk memboikot produk-produk yang terafiliasi dengan Israel. Salah satu akun yang menjadi sorotan dalam pembahasan ini adalah akun BDS Indonesia.

C. Pengumpulan Data

Pada tahap pengumpulan data, peneliti menggunakan library TweetHarvest untuk melakukan proses crawling data dari media sosial "X" (sebelumnya Twitter). TweetHarvest merupakan tools Python berbasis API yang memungkinkan pengambilan data secara otomatis dari akun atau kata kunci tertentu. Rentang waktu pengambilan data dilakukan dari 1 Januari 2024 hingga 31 Desember 2024, dengan batasan data hanya pada unggahan publik berbahasa Indonesia.

Dari proses crawling ini, berhasil dikumpulkan sebanyak 11.456 tweet mentah. Setelah melalui tahapan praproses (cleaning, normalisasi, dst.), diperoleh sebanyak 9.442 tweet bersih yang siap digunakan untuk tahap analisis dan pelabelan sentimen.

D. Kesimpulan

Langkah terakhir merupakan pembuatan laporan akhir, di mana kesimpulan dan saran akan dikemukakan setelah melalui serangkaian tahapan dalam analisis sentimen. Kesimpulan ini didasarkan pada hasil dari tahapan analisis sentimen yang telah dilakukan. Selain itu, laporan ini juga akan memberikan masukan yang dapat digunakan untuk pengembangan penelitian selanjutnya.

E. Diagram Blok

Pada Diagram Blok, Peneliti melakukan beberapa tahap penelitian dalam proses analisis sentimen yang ditunjukkan pada Gambar 2.



GAMBAR 2
DIAGRAM BLOK

F. Preprocessing

Tahapan preprocessing data dalam penelitian ini meliputi enam langkah utama. Pertama, cleaning data dilakukan dengan menghapus elemen yang tidak relevan seperti angka, tanda baca, mention, hashtag, URL, emotikon, karakter non-alfanumerik, dan spasi berlebih. Kedua, case folding mengubah seluruh huruf dalam teks menjadi huruf kecil agar konsisten. Ketiga, normalisasi diterapkan untuk mengganti bahasa informal menjadi bentuk baku guna meningkatkan kualitas analisis. Keempat, tokenizing memecah teks menjadi unit kata. Kelima, stopwords removal dilakukan untuk menghilangkan kata umum yang kurang bermakna seperti "dan", "di", dan "yang", dengan bantuan library Sastrawi. Terakhir, stemming digunakan untuk mengembalikan kata ke bentuk dasarnya, mengurangi variasi kata, dan menyederhanakan analisis.

G. Translate dan Labeling

Pada tahap pelabelan sentimen, peneliti menggunakan library TextBlob untuk mengklasifikasikan data ke dalam dua kategori, yaitu positif dan negatif. Sebelum dianalisis, teks yang semula berbahasa Indonesia diterjemahkan ke dalam bahasa Inggris menggunakan library Google Translate, karena TextBlob hanya mendukung analisis dalam bahasa Inggris. Proses ini diterapkan pada data yang telah melalui tahapan pra-pemrosesan sebelumnya untuk memastikan akurasi dalam identifikasi sentiment.

H. Visualisasi

Pada tahap ini, bertujuan untuk menyajikan hasil secara visual yang mudah dipahami dan menarik. Proses ini dilakukan dengan menggunakan library wordcloud untuk menampilkan representasi visual dari data teks. Selain itu, grafik juga akan ditampilkan dari hasil tahapan sebelumnya untuk memberikan gambaran yang lebih jelas.

I. Pembobotan Term (TF-IDF) dan Klasifikasi

Setelah pelabelan, data diproses melalui tahap seleksi fitur menggunakan metode TF-IDF untuk memberikan bobot lebih pada kata-kata yang memiliki pengaruh signifikan dalam klasifikasi. Representasi teks dengan TF-IDF membantu meningkatkan akurasi prediksi terhadap data baru. Tahap selanjutnya adalah klasifikasi, yang diimplementasikan menggunakan Python. Data dibagi menjadi dua bagian: data latih untuk membangun model dan data uji untuk mengevaluasi performa model. Pembagian ini bertujuan agar model dapat mempelajari pola dengan baik sekaligus menghindari overfitting, serta menguji kemampuan generalisasi terhadap data yang belum dikenali sebelumnya.

J. Evaluasi

Tahap akhir dalam proses ini adalah melakukan evaluasi terhadap model yang telah dibangun. Evaluasi dilakukan menggunakan metode confusion matrix untuk menilai tingkat akurasi sistem, serta menghitung metrik evaluasi lainnya seperti precision, recall, dan F1-score guna mengukur performa model Naïve Bayes secara lebih komprehensif.

IV. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap isu boikot produk yang terafiliasi dengan Israel melalui klasifikasi komentar pengguna di media sosial "X" ke dalam dua kategori, yaitu sentimen positif dan negatif. Proses penelitian meliputi tahapan pengumpulan data, praproses data, penerjemahan teks, pelabelan otomatis

menggunakan TextBlob, pembobotan fitur menggunakan TF-IDF, klasifikasi menggunakan algoritma Multinomial Naïve Bayes, serta evaluasi performa model berdasarkan metrik akurasi, presisi, recall, dan f1-score.

A. Pengumpulan dan Pra-Pemrosesan Data

Selama proses pengumpulan data, peneliti melaksanakan pengambilan postingan dari sosial media “X” mengaplikasikan library tweet-harvest yang ditunjukkan pada table berikut.

TABEL 1
PENGUMPULAN DATA

Created_At	Full_Text	Location
Mon Apr 29 17:12:35 +0000 2024	@aly_benzema @Israel Boikot produk zionis https://t.co/1ntT2N9QP3	Indonesia
Mon Apr 29 09:20:12 +0000 2024	@dr_koko28 Contoh IQ 58 mah liat hasil kelakuan HAMAS di GAZA termasuk kelakuannya menyalahgunakan bantuan pangan dan simpatisannya diberbagai negara yg pada bikin rusuh. Paling parah ide boikot produk2 Israel sedangkan rakyat GAZA yg dlm tekanan HAMAS hidup dari produk Israel.	Indonesia-West Java MC

Selanjutnya, komentar dikonversi menjadi teks bersih dengan tahapan cleaning, case folding, normalisasi, tokenizing, stopword removal, dan stemming. Tabel berikut menunjukkan hasil transformasi teks pada tiap tahap.

TABEL 2
TAHAP PREPROCESSING DATA

Tahap	Full_Text	Hasil
Cleaning	dimstrds hipohan kenapa netizen tidak murka dan boikot negara arab saudi yg jualan produk yg mendukung israel cuma mau tanya aja sambil memancing obrolan	dimstrds hipohan Kenapa netizen tidak murka dan boikot negara Arab Saudi yg jualan produk yg mendukung Israel Cuma mau tanya aja sambil memancing obrolan
Case Folding	dimstrds hipohan Kenapa netizen tidak murka dan boikot negara Arab Saudi yg jualan produk yg mendukung Israel Cuma mau tanya aja sambil memancing obrolan	dimstrds hipohan kenapa netizen tidak murka dan boikot negara arab saudi yg jualan produk yg mendukung israel cuma mau tanya aja sambil

		memancing obrolan
Normalisasi	dimstrds hipohan kenapa netizen tidak murka dan boikot negara arab saudi yg jualan produk yg mendukung israel cuma mau tanya aja sambil memancing obrolan	dimstrds hipohan kenapa netizen tidak murka dan boikot negara arab saudi yang jualan produk yang mendukung israel cuma mau tanya saja sambil memancing obrolan
Tokenasi	dimstrds hipohan kenapa netizen tidak murka dan boikot negara arab saudi yg jualan produk yg mendukung israel cuma mau tanya aja sambil memancing obrolan	['dimstrds', 'hipohan', 'kenapa', 'netizen', 'tidak', 'murka', 'dan', 'boikot', 'negara', 'arab', 'saudi', 'yang', 'jualan', 'produk', 'yang', 'mendukung', 'israel', 'cuma', 'mau', 'tanya', 'saja', 'sambil', 'memancing', 'obrolan']
Stopword Removal	['dimstrds', 'hipohan', 'kenapa', 'netizen', 'tidak', 'murka', 'dan', 'boikot', 'negara', 'arab', 'saudi', 'yang', 'jualan', 'produk', 'yang', 'mendukung', 'israel', 'cuma', 'mau', 'tanya', 'saja', 'sambil', 'memancing', 'obrolan']	dimstrds hipohan netizen murka boikot negara arab saudi jualan produk mendukung israel cuma mau tanya sambil memancing obrolan
Stemming	dimstrds hipohan netizen murka boikot negara arab saudi jualan produk mendukung israel cuma mau tanya sambil memancing obrolan	dimstrds hipohan netizen murka boikot negara arab saudi jual produk dukung israel cuma mau tanya sambil pancing obrol

B. Labeling dan Visualisasi

Setelah melalui tahap praproses meliputi penghapusan tanda baca, angka, tautan, serta normalisasi dan stemming, setiap komentar diterjemahkan dari Bahasa Indonesia penerjemahan ke dalam bahasa Inggris dijalankan dengan memanfaatkan Google Translate API. Langkah ini dianggap penting untuk mendukung kebutuhan analisis agar analisis sentimen dengan TextBlob dapat berjalan secara optimal, mengingat TextBlob dirancang untuk teks berbahasa Inggris. Proses pelabelan sentimen dilakukan secara otomatis, mengklasifikasikan komentar terbagi dalam dua kelas positif dan negatif. Dari total 9.442 komentar, sebanyak 63,1% (5.879 komentar) tergolong positif, sedangkan 36,9% (3.563 komentar) tergolong negatif.

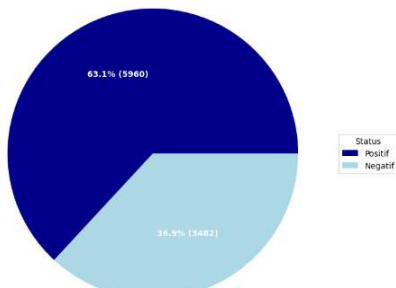
Stemming_Text	Sentimen
lebih gerai kfc malaysia tutup imbas gera boikot produk afiliasi israel lifestyle	Positif
drkoko contoh iq mah lihat hasil laku hamas gaza masuk laku menyalahgunakan bantu pangan simpatisan bagai negara pada bikin rusuh paling parah ide boikot produk israel rakyat gaza dalam tekan hamas hidup produk israel	Positif
dimstrds hipohan netizen murka boikot negara arab saudi jual produk dukung israel cuma mau tanya sambil pancing obrol	Negatif
deslini tetap semangat boikot produk afiliasi zionis israel freepalestine	Positif

Setelah melalui tahap preprocessing dan pelabelan sentimen, tahap selanjutnya adalah visualisasi data. Visualisasi diterapkan guna mempermudah pemahaman terhadap distribusi dan pola yang terkandung dalam dataset.



GAMBAR 3
VISUALISASI DATA DENGAN WORD CLOUD

WordCloud dipakai sebagai metode visualisasi guna memperlihatkan frekuensi kemunculan kata dalam data teks. Selain itu, digunakan juga diagram lingkaran (Pie Chart)



GAMBAR 4
VISUALISASI DATA DENGAN PIE CHART

untuk mengilustrasikan proporsi masing-masing kategori sentimen.

C. Pembobotan Term (TF-IDF) dan Klasifikasi Naïve

Bayes

Hasil pembobotan mengindikasikan bahwasanya meskipun suatu kata muncul satu kali dalam dokumen (TF = 1), bobot akhirnya berbeda tergantung nilai IDF-nya. Perkataan yang jarang muncul dalam dokumen lain akan bernilai Inverse Document Frequency (IDF) yang tinggi, sehingga menghasilkan bobot TF-IDF yang lebih besar. Sebaliknya, perkataan yang secara umum kerap muncul akan memiliki bobot yang lebih rendah dalam perhitungan TF-IDF.

Pos	Feature	TF	IDF	TF-IDF
(0, 45)	affiliation	1	5.6532	0.3150248997064443
(0, 362)	boys	1	3.8596	0.2150761800612109
(0, 516)	close	1	5.8223	0.3244465631432561
(0, 1271)	impact	1	5.2702	0.2936826232038828
(0, 1350)	israeli	1	2.6737	0.1489911727913409
(0, 1440)	kfc	1	3.6106	0.2011977562820183
(0, 1534)	lifestyle	1	8.3613	0.4659305058654740
(0, 1613)	malaysia	1	5.5992	0.3120119218132086
(0, 1755)	more	1	4.2076	0.2344683543915078
(0, 1801)	of	1	1.8054	0.1006077983443012

Pada proses klasifikasi, penelitian ini mengaplikasikan algoritma Multinomial Naïve Bayes sebagai metode utama, yang digabungkan dalam pipeline bersama teknik TF-IDF Vectorizer, dengan implementasi mengaplikasikan library Scikit-learn. TF-IDF diterapkan guna mentransformasikan teks hasil praproses (dalam kolom text_translate) menjadi representasi numerik berbasis bobot kata, yang kemudian menjadi input bagi algoritma Multinomial Naïve Bayes. Integrasi pipeline ini mempermudah proses pelatihan karena preprocessing dan klasifikasi dilakukan dalam satu alur otomatis.

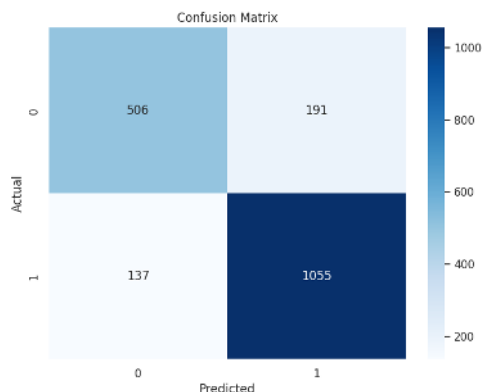
```
pipeline = Pipeline([
    ('tfidf', TfidfVectorizer()),
    ('nb', MultinomialNB())
])
```

Sebagaimana ditampilkan pada Gambar, kode pipeline menggabungkan tahap transformasi fitur dan pembelajaran model secara efisien.

Untuk mendapatkan performa model yang optimal, dilakukan pencarian kombinasi parameter terbaik mengaplikasikan metode Grid Search dengan skema 5-fold Cross-Validation. Parameter yang diuji mencakup konfigurasi ngram_range, min_df, max_df, dan sublinear_tf pada TF-IDF, serta alpha pada algoritma Naïve Bayes. Proses ini mengevaluasi berbagai kombinasi parameter untuk menentukan model dengan akurasi terbaik. Setelah mendapatkan model terbaik (best_estimator_), dilakukan evaluasi lanjutan guna mengkaji performanya terhadap data uji.

D. Evaluasi

Setelah tahap klasifikasi dilaksanakan dengan menerapkan algoritma Multinomial Naïve Bayes, tahap selanjutnya adalah melakukan evaluasi performa model guna mengevaluasi efektivitas dalam mengklasifikasikan komentar media sosial “X” ke dalam dua kategori sentimen, yakni positif (label 1) dan negatif (label 0). Penilaian ini dilakukan melalui confusion matrix yang mengilustrasikan performa model dengan menunjukkan berapa banyak prediksi yang tepat dan yang keliru.



Berdasarkan confusion matrix yang disajikan pada Gambar tersebut, model berhasil mengidentifikasi dengan benar sebanyak 506 data negatif (true negative) dan 1.055 data positif (true positive). Namun demikian, model juga melakukan kesalahan dengan mengklasifikasikan 191 data negatif menjadi positif (false positive) serta 137 data positif menjadi negatif (false negative).

	precision	recall	f1-score	support
0	0.79	0.73	0.76	697
1	0.85	0.89	0.87	1192
accuracy			0.83	1889
macro avg	0.82	0.81	0.81	1889
weighted avg	0.82	0.83	0.82	1889

Berdasarkan hasil evaluasi mengaplikasikan metrik klasifikasi, model memperlihatkan kinerja yang tergolong memuaskan dengan nilai akurasi sejumlah 83%. Nilai tersebut mengindikasikan bahwa mayoritas prediksi yang dihasilkan oleh model telah selaras dengan label yang sesungguhnya. Model menunjukkan kemampuan yang baik dalam mengenali komentar positif (kelas 1), sebagaimana terlihat dari nilai presisi sejumlah 85% dan recall sejumlah 89%.

V. KESIMPULAN

Berdasarkan hasil analisis terhadap 9.442 komentar, ditemukan bahwa mayoritas tanggapan bersentimen positif, yakni sebanyak 63,1% atau 5.958 komentar. Sementara itu, komentar dengan sentimen negatif berjumlah 3.484 atau sekitar 36,9%. Setelah data dibagi dengan perbandingan 80:20, didapat 7.553 data yang diaplikasikan untuk proses pelatihan dan 1.889 data untuk pengujian. Algoritma Naïve Bayes yang diterapkan sebagai model klasifikasi dalam studi ini mampu mencapai akurasi sejumlah 83%. Model tersebut lebih efektif dalam mengenali sentimen positif, yang direpresentasikan dengan nilai presisi sejumlah 85%, recall 89%, dan f1-score 87%. Sebaliknya, kemampuan model dalam mengidentifikasi sentimen negatif masih lebih rendah, dengan presisi 79%, recall 73%, dan f1-score 76%. Hasil ini mengindikasikan bahwa meskipun model cukup andal dalam mendeteksi komentar positif, masih terdapat beberapa komentar negatif yang salah diklasifikasikan sebagai positif.

Penelitian ini memberi kontribusi dalam memperluas wawasan masyarakat mengenai penerapan analisis sentimen berperan sebagai sarana untuk menelaah pandangan masyarakat terhadap berbagai permasalahan sosial, khususnya terkait gerakan boikot produk yang terafiliasi dengan Israel. Dengan memanfaatkan data komentar di

media sosial “X” dan menerapkan metode klasifikasi sentimen, penelitian ini menunjukkan bagaimana teknologi dapat diterapkan guna memetakan respons masyarakat secara sistematis. Diharapkan temuan dari penelitian ini tidak hanya menyajikan wawasan yang bermanfaat bagi masyarakat dalam mengevaluasi produk-produk yang dikaitkan dengan konflik politik dan kemanusiaan, tetapi juga berharga bagi merek-merek terkait untuk memahami persepsi publik, mengidentifikasi respons negatif, serta merancang pendekatan komunikasi yang lebih cermat dan fleksibel dalam merespons berbagai isu yang bersifat sensitif.

REFERENSI

- [1] M. Amri, “Konstruksi Realitas Konflik Israel-Palestina Di Media Online Republika.co.id dan Kompas.com,” *LayaR: Jurnal Ilmiah Seni Media Rekam*, 2021.
- [2] R. Sagita Dewi, “Pengaruh Konflik Palestina-Israel Terhadap Perekonomian Dunia,” *JEBESH: Journal of Economics Business Ethic and Science of History*, vol. 2, no. 3, 2024.
- [3] F. A. Ramadhan, “Peran Hukum Internasional dalam Menengahi Konflik Israel-Palestina pada Tahun 2023-2024,” *Rio Law Jurnal*, vol. 5, no. 1, 2023, doi: 10.36355/v1i2.
- [4] R. Sugandi and R. Anggraini, “Gerakan Sosial: Aksi Bela Palestina Boikot Produk Israel Di Kota Padang 2017-2023,” *CENDEKIA: Jurnal Ilmu Sosial, Bahasa dan Pendidikan*, vol. 4, no. 2, 2024.
- [5] Y. Nurtikasari, Syariful Alam, and Teguh Iman Hermanto, “Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naïve Bayes,” *INSOLOGI: Jurnal Sains dan Teknologi*, vol. 1, no. 4, pp. 411–423, Aug. 2022, doi: 10.55123/insologi.v1i4.770.
- [6] M. R. F. Septiazi and N. Yuliana, “Analisis Pengaruh Media Sosial Terhadap Gerakan Boikot Produk Israel Di Indonesia,” *Triwikrama: Jurnal Multidisiplin Ilmu Sosial*, vol. 2, no. 4, 2023.
- [7] A. Tiara Susilawati, N. Anjeni Lestari, and P. Alpria Nina, “Analisis Sentimen Publik Pada Twitter Terhadap Boikot Produk Israel Menggunakan Metode Naïve Bayes,” *Nian Tana Sikka: Jurnal Ilmiah Mahasiswa*, vol. 2, no. 1, pp. 26–35, 2024, doi: 10.59603/niantanasikka.v2i1.240.
- [8] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, “Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN,” *Jurnal KomtekInfo*, pp. 1–7, Jan. 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [9] T. A. Sari, E. Sinduningrum, and F. Noor Hasan, “Sentimen Ulasan Pelanggan Pada Aplikasi Fore Coffee Menggunakan Metode Naïve Bayes,” *KLIK: Kajian Ilmiah Informatika dan Komputer Analisis*, vol. 3, no. 6, pp. 773–779, 2023, doi: 10.30865/klik.v3i6.884.
- [10] B. Ramadhani and R. R. Suryono, “Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse,” *Jurnal Media Informatika Budidarma*, 2024, doi: 10.30865/mib.v8i2.7458.

- [11] P. Prihandoko, R. G. Alam, G. Gunawan, and D. Abdullah, *Memahami Konsep dan Implementasi Machine Learning*, 1st ed., vol. 1. Jambi: PT Sonpedia Publishing Indonesia, 2024.
- [12] Kusriani and E. T. Luthfi, *Algoritma Data Mining*. Yogyakarta: Penerbit Andi, 2009. Accessed: Jun. 29, 2025. [Online]. Available: <https://books.google.co.id/books?id=-Ojclag73O8C&printsec=frontcover&hl=id#v=onepage&q&f=false>
- [13] J. Khatib Sulaiman, F. Baehaqi, and N. Cahyono, "Analisis Sentimen Terhadap Cyberbullying Pada Komentar di Instagram Menggunakan Algoritma Naïve Bayes," *Indonesian Journal of Computer Science*, 2024.
- [14] R. A. Husen, R. Astuti, L. Marlia, R. Rahmadden, and L. Efrizoni, "Analisis Sentimen Opini Publik pada Twitter Terhadap Bank BSI Menggunakan Algoritma Machine Learning," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 211–218, Oct. 2023, doi: 10.57152/malcom.v3i2.901.
- [15] S. T. , M. Yessy Asri, M. K. Dr. Dra. Dwina Kuswardani, S. Kom. , M. K. Listra Frigia Missianes Horhoruw, and S. Siti Aisyah Ramadhana, *Analisis Sentimen Menggunakan Ulasan Pengguna Aplikasi*, 1st ed., vol. 1. Ponorogo: Uwais Inspirasi Indonesia, 2024. Accessed: Jun. 29, 2025. [Online]. Available: https://books.google.co.id/books/about/MACHINE_LEARNING_DEEP_LEARNING_Analisis.html?id=Yu7uEAAQBAJ&redir_esc=y