



Analisis Sentimen Pada X Dengan Hashtag #Indonesiagelap Menggunakan Metode Support Vector Machine (SVM)

Sentiment Analysis on X with the Hashtag #Indonesiagelap Using the Support Vector Machine (SVM) Method

Indah Mayang Sari Nipa¹, Rambu Yetti Kalaway² dan Reynaldy Thimotius Abineno³

Universitas Kristen Wira Wacana Sumba

Jl. R. Suprpto No.35, Prailiu, Kec. Kota Waingapu, Kabupaten Sumba Timur, Nusa Tenggara Timur

Corresponding author: mayangsariindah62@gmail.com, kalaway@unkriswina.ac.id

ABSTRACT

This study aims to analyze public opinion on the hashtag #IndonesiaGelap on social media platform X (Twitter) as a public response to various national issues following the 2025 presidential inauguration. Data was collected through a crawling technique using the tweet-harvest tool and analyzed using the Support Vector Machine (SVM) algorithm. After going through a preprocessing process (cleaning, normalization, tokenization, and TF-IDF), the data were categorized into two types of sentiment: positive and negative. The analysis results showed that the dominance of negative sentiment reached 69.58%. The SVM model showed an accuracy rate of 76.09%, with superior performance in identifying negative sentiment compared to positive sentiment. The F1-score for the negative class was 0.84, while for the positive class it was only 0.54. This indicates that although the model is generally quite accurate, there are still difficulties in identifying positive opinions. This study suggests data balancing, contextual labeling, and testing other methods for better results.

Keywords: *Sentiment Analysis, Platform X, Hashtag #indonesiagelap, Support Vector Machine (SVM).*

ABSTRAK

Penelitian ini ditujukan untuk menganalisis opini publik terhadap hashtag #IndonesiaGelap di platform media sosial X (Twitter) sebagai tanggapan masyarakat terhadap berbagai isu nasional setelah pelantikan presiden 2025. Data diambil melalui teknik crawling menggunakan perangkat tweet-harvest dan dianalisis dengan algoritma Support Vector Machine (SVM). Setelah melewati proses preprocessing (pembersihan, normalisasi, tokenisasi, dan TF-IDF), data dikategorikan ke dalam dua jenis sentimen: positif dan negatif. Hasil analisis menunjukkan bahwa dominasi sentimen negatif mencapai 69,58%. Setelah oversampling dilakukan, jumlah komentar pada kedua kelas menjadi sama, yaitu masing-masing 318 komentar atau 50% dari total data. Model SVM mencapai akurasi baik sebesar 90,62%, dengan kinerja tinggi dalam mengidentifikasi sentimen positif (F1-score 0,91) dan mendeteksi sentimen negatif (F1-score 0,90). Studi ini menyarankan penyeimbangan data, pelabelan kontekstual, dan pengujian metode lain demi hasil yang lebih baik.

Kata kunci: Analisis Sentimen, Platform X, Hashtag #indonesiagelap, Support Vector Machine (SVM).

PENDAHULUAN

Perkembangan teknologi dan internet, telah membawa perubahan dalam cara masyarakat berkomunikasi dan mendapatkan informasi (Lestari et al., 2024). Media sosial yang menjadi alat komunikasi orang-orang digunakan untuk mengemukakan pendapat atau opini dari berbagai kalangan untuk menyampaikan pendapat, opini maupun keluhan terhadap isu-isu yang berkembang. Salah satu platform media sosial yang paling populer adalah Twitter atau pada saat ini yang dikenal dengan nama X, platform memungkinkan pengguna menyampaikan opini,

keluhan, hingga dukungan secara real-time. di mana platform ini juga menjadi ruang diskusi tentang berbagai isu sosial, politik, dan ekonomi (Afandi, 2019).

Platform X menyediakan kemudahan penggunaan tagar (#), yang memudahkan riset dan penulisan topik. Di Indonesia, tagar sering digunakan untuk meningkatkan popularitas suatu topik dan menarik perhatian publik. Salah satu contohnya adalah munculnya #IndonesiaGelap, yang telah menjadi tren sejak 17 Februari 2025, sebagai bentuk kritik terhadap tindakan Presiden Prabowo Subianto selama 100 hari pertama pemerintahannya. Isu-isu utama yang dibahas adalah pembatasan kebebasan akademik, pembatasan subsidi, kontroversi seputar program Makanan Bergizi Gratis (MBG), dan dampak pendidikan terhadap kesehatan. Aksi Puncak terjadi pada 20 Februari 2025, dan diselesaikan dalam dua tahap pada 11 Maret 2025, sebagai respons terhadap Revisi UU TNI. Gerakan ini terbuka untuk mahasiswa dari berbagai daerah dan berlanjut hingga saat ini (Vindua & Zailani, 2023).

Dengan menggunakan tagar #IndonesiaGelap, pemerintah, termasuk Presiden Prabowo Subianto, menanggapi dengan pernyataan yang memicu beragam reaksi. Presiden menyatakan bahwa “Indonesia tampak cerah saat pagi hari” mengacu pada gerakan tersebut. Tanggapan ini memicu beragam respons dari publik di platform X, baik positif maupun negatif. Beberapa pengguna mengungkapkan sentimen negatif seperti “Semoga Indonesia menjadi lebih baik,” sementara yang lain mengkritik tindakan mahasiswa tersebut dan mendukung pemerintah. Ada juga yang mengkritik keras atau mempertanyakan validitas gerakan tersebut. Menurut pendapat ini, media sosial merupakan alat penting untuk menyalurkan aspirasi dan membangun komunikasi antara publik dan pemerintah, serta untuk menilai demokrasi digital di Indonesia (Fahrezi et al., 2024).

Fenomena digital seperti tagar #IndonesiaGelap menyoroti opini publik yang cepat dan meluas di media sosial. Untuk memahami ketahanan masyarakat, diperlukan analisis berbasis data, salah satunya melalui analisis sentimen (Styawati et al., 2021). Teknik ini digunakan untuk mengidentifikasi dan mengkategorikan opini pengguna, baik positif maupun negatif. Analisis sentimen penting tidak hanya untuk menilai respons terhadap kebijakan tetapi juga untuk membantu pemerintah dalam mengkaji kebijakan yang lebih responsif. Support Vector Machine (SVM) merupakan salah satu metode paling efektif dalam analisis ini karena dapat memproses input teks yang kompleks dan menghasilkan klasifikasi yang akurat (Darwis et al., 2020). Penelitian ini bertujuan untuk memahami sentimen publik terhadap isu-isu dalam tagar #IndonesiaGelap dengan menggunakan metode SVM dalam analisis sentimen media sosial.

Diharapkan penelitian ini akan memberikan kontribusi untuk memahami sentimen publik tentang isu-isu penting di Indonesia dan menawarkan rekomendasi kepada organisasi-organisasi terkait untuk pengambilan keputusan yang lebih baik berdasarkan data yang tersedia.

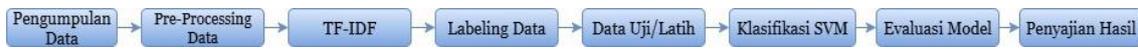
MATERI DAN METODE

X adalah alat media sosial realtime yang berkembang cukup pesat di era modern ini. X menjadi media sosial yang banyak dimanfaatkan oleh remaja hingga orangtua khususnya mahasiswa (Donoriyanto et al., 2023). Hashtag mempermudah netizen untuk membahas topik yang sedang hangat dan menjadi bagian dari trending topic (Fauziah, 2025).

Analisis sentimen merupakan bidang penelitian yang menganalisis sentimen, pendapat, evaluasi, sikap dan emosi orang dari bahasa tertulis. analisis sentimen juga bisa diartikan sebagai studi komputasi tentang penilaian, sikap, pendapat dan emosi orang terhadap entitas, individu, peristiwa, masalah, topik, serta atributnya (Pebrianti, 2025).

Algoritma SVM merupakan metode pembelajaran terawasi yang menghasilkan pemetaan dari fungsi input-output dari sekumpulan data training. Fungsi pemetaan ini berupa fungsi-fungsi klasifikasi atau fungsi regresi (Yulia Ardana et al., 2023).

Untuk menjelaskan tahapan penelitian secara sistematis, berikut disajikan alur proses analisis sentimen menggunakan metode Support Vector Machine (SVM) seperti ditunjukkan pada Gambar berikut.



Gambar 1. Alur Penelitian

Gambar di atas menunjukkan alur proses analisis sentimen menggunakan metode Support Vector Machine (SVM). Proses dimulai dengan pengumpulan data dari media sosial, dilanjutkan dengan pre-processing yang meliputi tahap cleaning, folding case, normalisasi, tokenisasi, stopword removal, dan stemming untuk mempersiapkan data teks. Selanjutnya, data diubah menjadi representasi numerik menggunakan metode TF-IDF, lalu dilakukan pelabelan data berdasarkan sentimen. Setelah itu, data dibagi menjadi data latih dan data uji, yang masing-masing digunakan dalam proses training dan testing. Model SVM kemudian digunakan untuk melakukan klasifikasi sentimen, dan tahap akhir adalah penyajian hasil dalam bentuk visualisasi dan evaluasi performa model.

HASIL DAN PEMBAHASAN

PENGUMPULAN DATA

Penelitian ini mengumpulkan data dari media sosial X melalui proses crawling menggunakan alat tweet-harvest, yang memungkinkan pengambilan tweet berdasarkan kata kunci tanpa akses API resmi. Fokus pengambilan data adalah pada hashtag #IndonesiaGelap, dalam rentang waktu Februari hingga Juni 2025, yaitu periode munculnya berbagai isu penting pasca pelantikan presiden baru. Proses crawling dilakukan melalui perintah baris di Google Colab dan hasilnya disimpan dalam format CSV. Setiap tweet yang dikumpulkan memuat informasi seperti waktu unggahan, nama pengguna, isi tweet, jumlah retweet dan like, serta tautan media jika tersedia. Data tersebut kemudian disiapkan untuk tahap selanjutnya, yakni preprocessing, pelabelan sentimen, hingga klasifikasi, yang menjadi inti dari proses analisis sentimen dalam penelitian ini.

	A	B	C	D	E	F	G	H
1	conversation_id	created_at	favorite_count	full_text	id_str	image_url	in_reply_to_screen_name	loc
2	1935541465884	Thu Jun 19 03:3	0	bos buzzer yang menggoreng semu media mulai dari isu RUU TNI Indonesia Gelap Kejaksan Agung sampai PEMERINTAH	1935541465884	https://pbs.twimg.com/media/Gt...	in	
3	193540902275	Thu Jun 19 03:3	0	Marcella Santoso dalam dibalik kehebohan kemarin! Dia udah memberikan pernyataan langsung dari #MarcellaSantoso #Prab	193540902275	https://pbs.twimg.com/media/Gt...	in	
4	1935405251399	Thu Jun 19 03:2	1	@MasBRO_back @DarsAlexandra1 @dewapers @kemkomdigi Satu lagi bukti betapa sungguh Mulyondol punya kepantasan	1935405251399	https://pbs.twimg.com/media/Gt...	in	
5	1935539910974	Thu Jun 19 03:2	0	Apa cerita kasus Oplis Pertamina ini? Bagaimana kelanjutan hukumnya wahai BANGSA BESAR ?? UMKM Primogem #Iran #	1935539910974	https://pbs.twimg.com/media/Gt...	in	
6	1935316566419	Thu Jun 19 03:2	0	@tempodotco Pengangguran bukan persoalan ada atau tidaknya job fair tapi soal pemerintah Prabowo gagal membangun ekor	1935316566419	https://pbs.twimg.com/media/Gt...	in	
7	1935537330678	Thu Jun 19 03:1	0	Aku pribadi sih masih tertangis ngiang yg bilang Apa pula #IndonesiaGelap Muka kau yg gelap Sumpah puncak komedi	1935537330678	https://pbs.twimg.com/media/Gt...	in	
8	1935535812218	Thu Jun 19 03:1	4	BUZZER TOM LEMBONG! Kejagung tetapkan tersangka pengacara Marcella Santoso dan Junaedi Saibih serta Dir. JakTV Tia	1935535812218	https://pbs.twimg.com/media/Gt...	in	
9	1935533310945	Thu Jun 19 03:0	4	Kejagung tetapkan tersangka pengacara Marcella Santoso dan Junaedi Saibih serta Dir. JakTV Tian Bahtiar. Kotiganya	1935533310945	https://pbs.twimg.com/media/Gt...	in	
10	1935531181984	Thu Jun 19 02:5	0	Marcella Santoso ngaku dalam hoaks RUU TNI Indonesia Gelap sampe nyerang Kejagung & Prabowo! Yang kemarin	1935531181984	https://pbs.twimg.com/text_...	in	
11	1935530888345	Thu Jun 19 02:5	0	Marcella Santoso ngaku dalam hoaks RUU TNI Indonesia Gelap sampe nyerang Kejagung & Prabowo! Yang kemarin	1935530888345	https://pbs.twimg.com/media/Gt...	in	
12	1935530864399	Thu Jun 19 02:5	0	Marcella Santoso ngaku dalam hoaks RUU TNI Indonesia Gelap sampe nyerang Kejagung & Prabowo! Yang kemarin	1935530864399	https://pbs.twimg.com/media/Gt...	in	
13	1935530821487	Thu Jun 19 02:5	0	Marcella Santoso ngaku dalam hoaks RUU TNI Indonesia Gelap sampe nyerang Kejagung & Prabowo! Yang kemarin	1935530821487	https://pbs.twimg.com/media/Gt...	in	
14	1935524216184	Thu Jun 19 02:2	0	Bos Buzzer yang Goreng Isu di Negara #MarcellaSantoso #Prabowo #Korupsi #KasusMK #KasusGula #KasusCPO #Buzzer	1935524216184	https://pbs.twimg.com/media/Gt...	in	
15	1935268751307	Thu Jun 19 02:2	1	@tempodotco...lka orang tolot berkumpul n bikin UU namanya LUJ goblok lalu dianulir MK kok Rakyat disalahkan mirip	1935268751307	https://pbs.twimg.com/media/Gt...	in	
16	1935189884566	Thu Jun 19 02:1	0	@jhoristorus_19 Kira? @jhoristorus_19 bakal keseret ga oleh pengakuan Marcella Santoso? Siapa pun yg menerima	1935189884566	https://pbs.twimg.com/media/Gt...	in	
17	1935505397588	Thu Jun 19 02:1	0	@DokterTifa @DanSandiU adalah contoh dari korban virus tpu-tolot tingkat kronis. Sebenarnya kasihan juga hanya	1935505397588	https://pbs.twimg.com/media/Gt...	in	
18	1935502060742	Thu Jun 19 01:5	1	@ch_chotimah2 Parah... parah pejabat sudah ga pada bener. sampai? Haji saja bermasalah. yg bener apa coba	1935502060742	https://pbs.twimg.com/media/Gt...	in	
19	1935312771304	Thu Jun 19 01:5	0	@btdrbekasi Refleksi PEMERINTAH RUSATI #IndonesiaGelap	1935312771304	https://pbs.twimg.com/media/Gt...	in	
20	1935513453130	Thu Jun 19 01:4	3	Dari RUU TNI #IndonesiaGelap sampai fitnah ke Pemerintahan #Prabowo semuanya digoreng lim buzzer Marcella	1935513453130	https://pbs.twimg.com/media/Gt...	in	
21	1935309462114	Thu Jun 19 01:4	0	@DarsAlexandra1 @Aprilia.in adalah salah satu produk masalah akibat virus tpu-tolot yang merajalela. Kar	1935309462114	https://pbs.twimg.com/media/Gt...	in	
22	1935511887304	Thu Jun 19 01:3	56	seperjangan mesh ada gorengan #IndonesiaGelap hanya bikin2an saja. https://t.co/UY3npljvS	1935511887304	https://pbs.twimg.com/media/Gt...	in	
23	1935185471161	Thu Jun 19 01:2	0	@tempodotco Maktum hukumnya bisa dibeli #IndonesiaGelap	1935185471161	https://pbs.twimg.com/media/Gt...	in	
24	1935329148877	Thu Jun 19 01:2	0	@tempodotco Efek domino korupsi #IndonesiaGelap	1935329148877	https://pbs.twimg.com/media/Gt...	in	
25	1935309112470	Thu Jun 19 01:2	0	@NerikMonica Maktum mmmomom belum mati #IndonesiaGelap	1935309112470	https://pbs.twimg.com/media/Gt...	in	

Gambar 2. Hasil Crawling Data

PREPROCESSING DATA

Langkah selanjutnya setelah pengumpulan data adalah prapemrosesan, yang juga dikenal sebagai pembersihan data. Langkah ini krusial untuk mengubah teks menjadi format terstruktur yang dapat dipahami oleh algoritma pembelajaran mesin. Prapemrosesan dilakukan dengan beberapa cara:

1. Cleaning

Pada tahap awal preprocessing, dilakukan proses cleaning untuk membersihkan data dari elemen-elemen yang tidak relevan, seperti simbol, karakter khusus, mention, hashtag, emoji, dan tanda baca. Hal ini penting karena data dari media sosial seperti X cenderung tidak terstruktur. Salah satu langkah pembersihan melibatkan penghapusan emoji dan simbol Unicode menggunakan fungsi Python bernama `remove_emoji()`, yang memanfaatkan ekspresi reguler (regex) untuk mengenali dan menghapus karakter-karakter tersebut. Fungsi ini dirancang untuk memastikan input valid dan kemudian membersihkan teks dari berbagai simbol, seperti ikon transportasi, bendera, dan bentuk geometris. Dengan menghilangkan elemen tersebut, data menjadi lebih bersih dan siap diproses pada tahap selanjutnya, seperti tokenisasi, stemming, dan transformasi TF-IDF dalam analisis sentimen.

Tabel 1. Hasil Cleaning

Sebelum	Sesudah
@@andreasharsono Resiko bakal terulang lg gaes #CabutUUTNI #TolakRUUPolri #IndonesiaGelap	Resiko bakal terulang lg gaes CabutUUTNI TolakRUUPolri IndonesiaGelap

2. Folding Case

Setelah proses cleaning, tahap berikutnya adalah case folding, yaitu mengubah seluruh huruf dalam teks menjadi huruf kecil untuk menyeragamkan penulisan. Proses ini mencegah perbedaan interpretasi kata akibat variasi kapitalisasi, seperti "Indonesia", "INDONESIA", dan "indonesia" yang disamakan menjadi "indonesia". Dalam penelitian ini, digunakan fungsi `case_folding()` yang memeriksa input bertipe string dan mengubahnya menggunakan `text.lower()`. Hasilnya, seluruh huruf kapital berhasil dikonversi, sehingga kata seperti "CabutUUTNI" dan "#IndonesiaGelap" menjadi "cabutuutni" dan "indonesiagelap". Penyeragaman ini penting agar model tidak menganggap kata yang sama sebagai entitas berbeda hanya karena perbedaan kapitalisasi.

Tabel 2. Hasil Folding Case

Sebelum	Sesudah
Resiko bakal terulang lg gaes CabutUUTNI TolakRUUPolri IndonesiaGelap	resiko bakal terulang lg gaes cabutuutni tolakruupolri indonesiagelap

3. Normalisasi

Normalisasi merupakan salah satu tahapan penting dalam preprocessing data teks yang bertujuan untuk menyamakan bentuk kata agar sesuai dengan kaidah penulisan yang baku. Pada script di atas, proses normalisasi dilakukan dengan mengganti kata-kata tidak baku menjadi bentuk bakunya. Script tersebut diawali dengan komentar #terapkan fungsi penggantian kata tidak baku yang menjelaskan bahwa fungsi yang digunakan bertujuan untuk menangani kata-kata yang tidak sesuai dengan ejaan yang disempurnakan (EYD).

Tabel 3. Normalisasi

Sebelum	Sesudah
resiko bakal terulang lg gaes cabutuutni tolakruupolri indonesiagelap	resiko bakal terulang lagi gaes cabutuutni tolakruupolri indonesiagelap

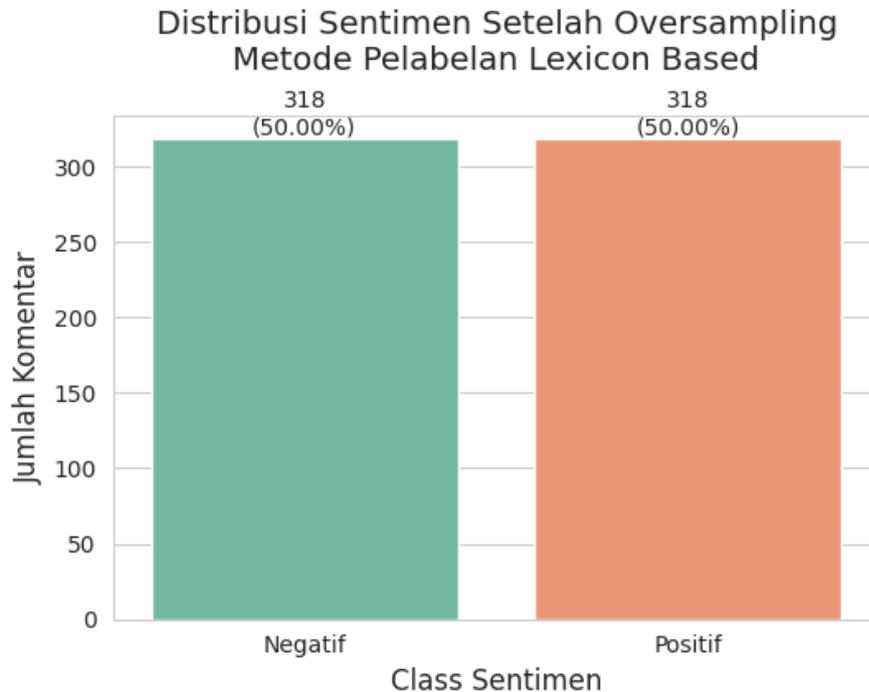
4. Tokenize

Pada tahap tokenisasi ini, proses konversi teks atau data menjadi sejumlah kecil unit yang dikenal sebagai token akan dilakukan. Token ini merupakan langkah yang sangat penting karena berfungsi sebagai fondasi untuk langkah-langkah analisis selanjutnya, seperti stopword dan stemming.

Tabel 4. Hasil Tokenize

Sebelum	Sesudah
---------	---------

Sebaliknya, terdapat 139 komentar dengan sentimen positif, atau sekitar 30,4%. Hasil tersebut menunjukkan bahwa mayoritas pengguna atau masyarakat umum memiliki opini negatif terhadap topik yang sedang dibahas. Tingginya komentar negatif dapat menjadi tanda kurangnya antusiasme, kritik, atau pemahaman terhadap topik yang sedang dibahas. Oleh karena itu, perlu untuk lebih memperhatikan faktor-faktor yang berkontribusi terhadap respons negatif ini agar dapat merespons secara tepat dan efektif.

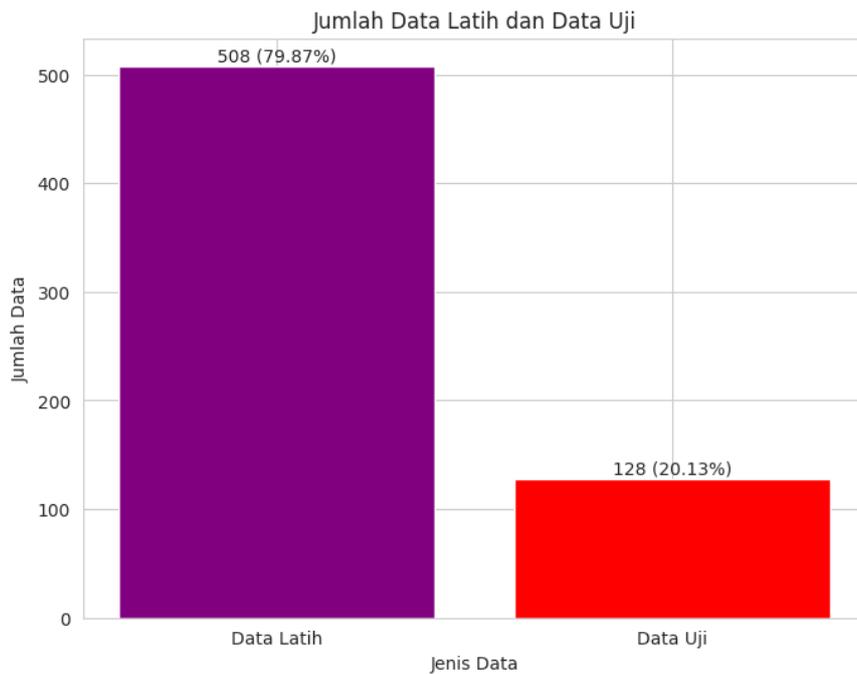


Gambar 5. Grafik Setelah Oversampling

Sebagai langkah untuk mengatasi ketidakseimbangan data tersebut, dilakukan proses oversampling yang bertujuan untuk meningkatkan jumlah data pada kelas minoritas (sentimen positif) sehingga distribusi antara kelas positif dan negatif menjadi seimbang. Setelah oversampling dilakukan, jumlah komentar pada kedua kelas menjadi sama, yaitu masing-masing 318 komentar atau 50% dari total data. Dengan distribusi data yang lebih seimbang ini, model klasifikasi yang akan dibangun diharapkan dapat belajar secara lebih adil dari kedua jenis sentimen dan menghasilkan prediksi yang lebih akurat. Selain itu, pendekatan ini juga membantu mencegah bias model terhadap kelas mayoritas, yang dalam hal ini adalah sentimen negatif. Ke depannya, hasil analisis yang lebih objektif dapat digunakan untuk merancang strategi komunikasi yang lebih efektif dan responsif terhadap opini publik.

DATA UJI/LATIH

Setelah proses prapemrosesan dan sentimen selesai, langkah selanjutnya adalah menganalisis data menjadi dua bagian utama, yaitu data latih dan uji. Dari 457 titik data yang tersedia, 365% data latih dan 92% data uji digunakan dalam studi ini. Tujuan studi ini adalah untuk mendeteksi overfitting dan memastikan bahwa model dapat menggeneralisasi informasi baru secara akurat. Dengan demikian, proses pelatihan model lebih efektif karena menggunakan data yang independen dari data latih.



Gambar 6. Grafik Pembagian Data Latih dan Data Uji

KLASIFIKASI SVM

Langkah selanjutnya adalah pembuatan dan pelatihan model klasifikasi menggunakan algoritma Support Vector Machine (SVM). Dalam hal ini, fungsi kernel linear digunakan untuk menganalisis data berdasarkan grafik atau hiperbidang. Setelah proses ini selesai, model digunakan untuk membuat prediksi tentang data baru berdasarkan pembelajaran yang diperoleh dari data tersebut. Model kemudian diimplementasikan dalam variabel SVM dan dilatih menggunakan data siswa yang telah dikonversi menjadi vektor.

```
#MEMBUAT MODEL SVM
svm = SVC(kernel='linear')
svm.fit(X_train_vectorized, y_train)
```

SVC

SVC(kernel='linear')

Gambar 7. Klasifikasi SVM

EVALUASI MODEL

Berikut menunjukkan hasil evaluasi model SVM menunjukkan akurasi sebesar 90,62%, yang berarti sekitar 91 dari 100 prediksi yang dihasilkan oleh model sesuai dengan label sebenarnya. Berdasarkan laporan klasifikasi, model menunjukkan kinerja yang sangat baik dan seimbang pada kedua kelas. Untuk kelas Negatif, nilai presisi tercatat sebesar 0,89, recall sebesar 0,92, dan skor f1 sebesar 0,90. Sementara itu, pada kelas Positif, model bahkan memiliki presisi yang lebih tinggi yaitu 0,92, recall sebesar 0,90, dan skor f1 sebesar 0,91. Nilai rata-rata makro dan rata-rata tertimbang untuk presisi, recall, dan f1-score semuanya sebesar 0,91. Hal ini mengindikasikan bahwa model mampu mengenali kedua kelas dengan baik tanpa menunjukkan bias terhadap salah satu kelas.

```

SVM Accuracy: 0.90625
SVM Accuracy: 90.62%

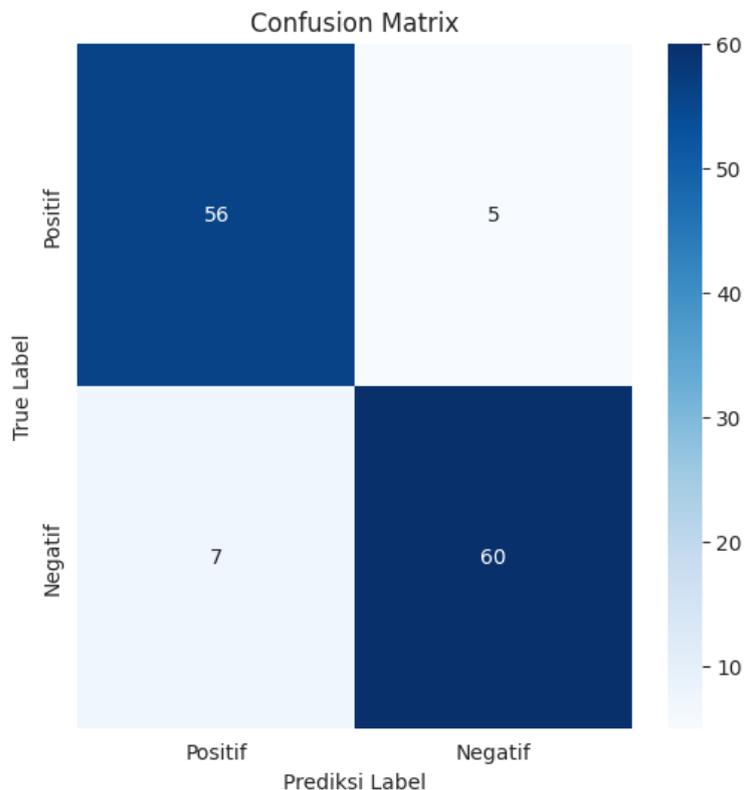
SVM Classification Report:

```

	precision	recall	f1-score	support
Negatif	0.89	0.92	0.90	61
Positif	0.92	0.90	0.91	67
accuracy			0.91	128
macro avg	0.91	0.91	0.91	128
weighted avg	0.91	0.91	0.91	128

Gambar 8. Hasil Akurasi, Presisi, Recall, dan F1-score

Terdapat matriks konfusi pada langkah berikutnya, yang menunjukkan kinerja klasifikasi model untuk dua kelas, "positif" dan "negatif". Dari semua data uji, sekitar 56 titik data positif berhasil diklasifikasikan, sementara tiga titik data positif diklasifikasikan sebagai data negatif. Di sisi lain, hanya 60 titik data yang diklasifikasikan sebagai negatif untuk kelas negatif, sementara 7 titik data diklasifikasikan sebagai positif. Hasilnya menunjukkan bahwa model SVM berkinerja lebih baik ketika menganalisis data positif daripada data negatif, tetapi masih memiliki keterbatasan ketika menganalisis data negatif yang sering diklasifikasikan sebagai positif. Hal ini juga selaras dengan nilai recall dan f1-score kelas negatif yang lebih rendah.

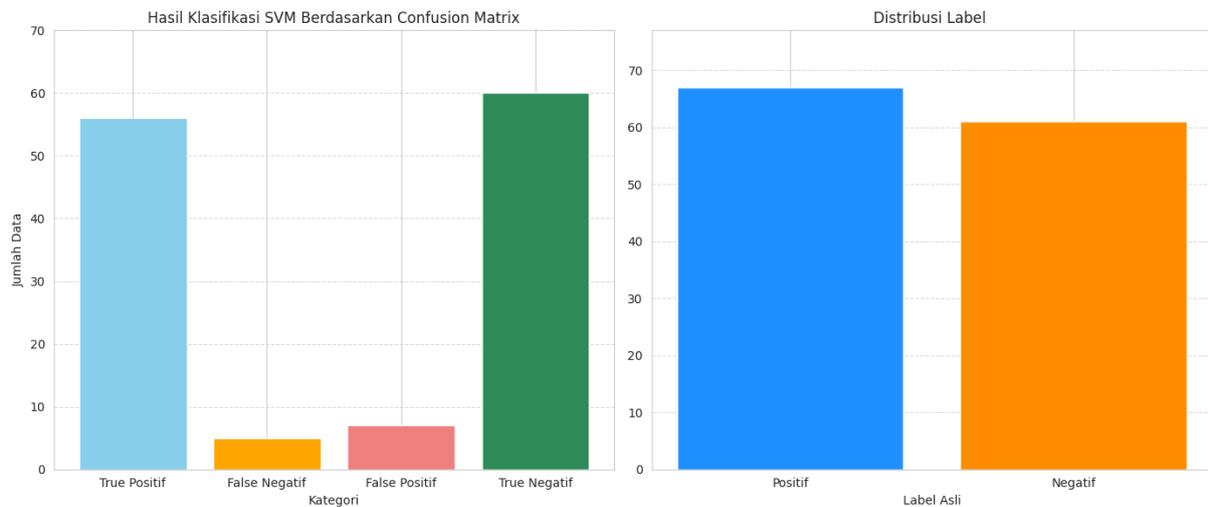


Gambar 9. Hasil Confusion Matrix

PENYAJIAN HASIL

Support Vector Machine (SVM), model klasifikasi yang digunakan dalam studi ini, dievaluasi menggunakan dua jenis grafik yang berbeda. Grafik pertama menampilkan hasil

klasifikasi model berdasarkan matriks konfusi, sementara grafik kedua menampilkan distribusi label aktual dari data yang digunakan. Kedua grafik ini memberikan ilustrasi komprehensif tentang efektivitas model dalam menganalisis sentimen dan karakteristik dari data yang dianalisis.



Gambar 10. Grafik Penyajian Hasil

Grafik di atas mengilustrasikan kinerja model SVM dan distribusi label data uji. Grafik kiri menampilkan hasil klasifikasi berdasarkan matriks kebingungan, dengan True Negatif mendominasi (60 data), yang menunjukkan model yang efektif untuk menganalisis sentimen negatif. Total sekitar 56 nilai True Positive dan False Negative menunjukkan bahwa model tidak berhasil menyajikan data positif. False Positive tercatat 7 menunjukkan masalah dengan kelas relatif negatif yang kecil.

Grafik kanan mengilustrasikan distribusi label asli, dengan data negatif (61), lebih banyak daripada data positif (67). Ketidakseimbangan ini memengaruhi kinerja model, yang lebih akurat untuk kelas mayoritas negatif daripada yang positif. Kedua grafik ini menggambarkan kesulitan dalam klasifikasi akibat distribusi data yang tidak merata.

KESIMPULAN

Penelitian ini menganalisis sentimen publik terhadap hashtag #IndonesiaGelap di platform X sebagai respons terhadap kebijakan pemerintah awal masa jabatan Presiden 2025. Dengan menggunakan metode Support Vector Machine (SVM), data dianalisis setelah melalui tahapan preprocessing, TF-IDF, dan pelabelan berbasis lexicon. Hasil menunjukkan bahwa sentimen negatif mendominasi tanggapan masyarakat sebesar 69,58%, mencerminkan keresahan terhadap isu-isu seperti pengurangan subsidi, revisi UU TNI, dan program Makanan Bergizi Gratis. Setelah oversampling dilakukan, jumlah komentar pada kedua kelas menjadi sama, yaitu masing-masing 318 komentar atau 50% dari total data. Model SVM mencapai akurasi baik sebesar 90,62%, dengan kinerja tinggi dalam mengidentifikasi sentimen positif (F1-score 0,91) dan mendeteksi sentimen negatif (F1-score 0,90).

Untuk penelitian selanjutnya, disarankan melakukan penyeimbangan data antar kelas, mempertimbangkan pelabelan manual berbasis konteks, serta mengeksplorasi algoritma lain seperti Naive Bayes, Random Forest, atau deep learning (misalnya LSTM dan BERT). Penambahan fitur seperti emotikon, hashtag, dan bahasa informal juga dapat meningkatkan akurasi klasifikasi. Selain itu, perluasan jumlah data dan periode pengambilan akan memberikan pemahaman yang lebih komprehensif terhadap dinamika opini publik di media sosial.

DAFTAR PUSTAKA

- Afandi, Y. (2019). Gereja dan Pengaruh Teknologi Informasi “Digital Ecclesiology.” *Fidei: Jurnal Teologi Sistematis Dan Praktika*, 1(2), 270–283. DOI: <https://doi.org/10.34081/fidei.v1i2.12>
- Darwis, D., Pratiwi, E. S., & Pasaribu, A. F. O. (2020). Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. *Edutic - Scientific Journal of Informatics Education*, 7(1), 1–11. DOI: <https://doi.org/10.21107/edutic.v7i1.8779>
- Donoriyanto, D. S., Indiyanto, R., Juliardi A. R., N. R., & Syamsiah, Y. A. (2023). Optimalisasi Penggunaan Media Sosial sebagai Sarana Promosi Online Store pada Pelaku UMKM di Kota X. *Jurnal Abdimas Peradaban*, 4(1), 42–50.
- Fahrezi, I. A., Rudiman, & Nauval Azmi Verdikha. (2024). Analisis Sentimen Twitter Atas Isu Hak Angket Menggunakan Pembobotan TF-IDF dan Algoritma SVM. *Sci-Tech Journal*, 3(2), 179–192. DOI: <https://doi.org/10.56709/stj.v3i2.526>
- Fauziah, R. (2025). Eksistensi Relawan Digital Dalam Penyebaran Hashtag #Prayforrua Di “X” Sebagai Respon Pada Manajemen Bencana. *Journal of Science and Social Research*, 4307(1), 421–429. DOI: <https://doi.org/10.54314/jssr.v8i1.2735>
- Lestari, E., Fatonah, U., Syifa, N., Rahmadayanti, N., Azizah, N., Nur Faidah, A., Author, C., Studi Manajemen, P., Lambung Mangkurat, U., Brigadir Jenderal Hasan Basry, J. H., & Selatan, K. (2024). Sosialisasi Dampak Perkembangan Teknologi dan Internet pada Generasi Z di Banjarmasin. *Damhil: Jurnal Pengabdian Pada Masyarakat*, 3(2), 82–93.
- Pebrianti, R. D. (2025). Analisis Sentimen Masyarakat Platform X Terhadap Korupsi Pt. Pertamina (Persero) Menggunakan Metode Svm. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(2). DOI: <https://doi.org/10.23960/jitet.v13i2.6399>
- Styawati, S., Hendrastuty, N., & Isnain, A. R. (2021). Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine. *Jurnal Informatika: Jurnal Pengembangan IT*, 6(3), 150–155. DOI: <https://doi.org/10.30591/jpit.v6i3.2870>
- Vindua, R., & Zailani, A. U. (2023). Analisis Sentimen Pemilu Indonesia Tahun 2024 Dari Media Sosial Twitter Menggunakan Python. *JURIKOM (Jurnal Riset Komputer)*, 10(2), 479.
- Yulia Ardana, Ridwan A. Kambau, & Mustikasari. (2023). Perbandingan Analisis Sentimen Algoritma Support Vector Machine Dan Naïve Bayes Terhadap Tanggapan Publik Tentang Pembelajaran Online Di Masa Pandemi Covid-19. *AGENTS: Journal of Artificial Intelligence and Data Science*, 3(1), 9–15. DOI: <https://doi.org/10.24252/jagti.v3i1.46>