

Opini Publik Pasca-Pemilihan Presiden: Eksplorasi Analisis Sentimen Media Sosial X Menggunakan SVM

Khoirul Adib¹, Maya Rini Handayani², Wenty Dwi Yuniarti³, Khotibul Umam⁴

^{1,2,3,4} Teknologi Informasi, Fakultas Sains dan Teknologi, UIN Walisongo Semarang, Indonesia
Jl. Prof. Dr. Hamka, Semarang 50185, Jawa Tengah, Indonesia

2108096052@student.walisongo.ac.id¹, maya@walisongo.ac.id², wenty@walisongo.ac.id³,
Khotibul_umam@walisongo.ac.id⁴,

Received : April, 2024

Accepted : May, 2024

Published : August, 2024

Abstract

Presidential elections in Indonesia often serve as triggers for dramatic changes in public opinion dynamics, especially in the digital era filled with voices scattered across social media. This research aims to map changes in public sentiment post-presidential elections using social media analysis, focusing on application X, which has 24 million active users in Indonesia. The Support Vector Machine (SVM) method is utilized to analyze and classify sentiments accurately based on trending tweet words after the presidential elections. This study aims to provide a deeper understanding of post-presidential election public opinion changes by depicting the dynamics of societal sentiments reflected in social media. The contribution of this research lies in accurately mapping changes in public opinion, which can provide valuable insights for policymakers, political analysts, and social media practitioners in responding to the needs of society in this digital era. Testing results using a dataset of 3850 tweets with three trending tweet word classes from platform X showed the highest accuracy rates in classifying "Peaceful Election" at 97.3%, "Inquiry Rights" at 96.5%, and "Election Fraud" at 94.0%.

Keywords: Sentiment Analysis, Post-Presidential Election, Support Vector Machine (SVM)

Abstrak

Pemilihan Presiden di Indonesia seringkali menjadi pemicu perubahan dramatis dalam dinamika opini publik, terutama di era digital yang dipenuhi dengan suara yang tersebar di media sosial. Penelitian ini bertujuan untuk memetakan perubahan sentimen publik pasca-pemilihan Presiden dengan menggunakan analisis media sosial, dengan fokus pada aplikasi X yang memiliki 24 juta pengguna aktif di Indonesia. Metode Support Vector Machine (SVM) digunakan untuk menganalisis dan mengklasifikasikan sentimen dengan akurat berdasarkan kata tweet yang sedang tren setelah pemilihan Presiden. Penelitian ini bertujuan untuk memberikan pemahaman yang lebih dalam tentang perubahan opini publik pasca-pemilihan presiden, dengan menggambarkan dinamika sentimen masyarakat yang tercermin dalam media sosial. Kontribusi dari penelitian ini adalah pemetaan yang akurat tentang perubahan opini publik, yang dapat memberikan wawasan yang berharga bagi pembuat kebijakan, analisis politik, dan praktisi media sosial dalam merespons kebutuhan masyarakat di era digital ini. Hasil pengujian dengan menggunakan 3850 dengan karakteristik dataset dengan menggunakan tiga kelas kata tweet yang sedang tren dari platform X menunjukkan tingkat akurasi tertinggi pada klasifikasi "Pemilu Damai" dengan 97.3%, "Hak Angket" dengan 96.5%, dan "Pemilu Curang" dengan 94.0%.

Kata Kunci: Analisis Sentimen, Pasca-Pemilihan Presiden, Support Vector Machine (SVM)

1. PENDAHULUAN

Pemilihan Presiden merupakan salah satu momen penting dalam demokrasi yang memberikan kesempatan kepada warga negara untuk menentukan arah dan masa depan negara mereka. Khususnya di Indonesia, pemilihan Presiden ini diadakan sebagai pesta rakyat lima tahunan yang sudah disahkan dalam Undang-Undang No. 42 Tahun 2008[1]. Proses pemilihan tidak hanya mencerminkan prinsip dasar kedaulatan rakyat, tetapi juga berfungsi sebagai sarana evaluasi dan perubahan dalam pemerintahan. Pemilihan Presiden sering kali diikuti oleh debat publik yang intens, kampanye politik, dan analisis mendalam tentang visi, misi, serta kebijakan yang ditawarkan oleh para kandidat.

Di era digital saat ini, media sosial telah menjadi platform utama bagi masyarakat untuk menyampaikan opini, gagasan, dan perasaan mereka tentang berbagai topik, termasuk politik. Khususnya pada Pemilihan Presiden tahun ini, sebagai salah satu event politik paling signifikan, secara khusus menarik perhatian dan diskusi publik yang luas. Melihat ritme pada platform X media sosial yang global dan dinamis[2], ini telah menjadi salah satu sarana utama bagi masyarakat untuk bertukar gagasan, opini, dan informasi tentang berbagai topik, termasuk pemilihan Presiden[3]. Berdasarkan laporan *We Are Social dan Hootsuite* pada tahun 2024 menyebutkan bahwa Platform X memiliki pengguna sekitar 24 juta di Indonesia[4], perlu diketahui media social ini kemampuannya untuk menyebarkan informasi secara cepat dan luas membuat X menjadi alat yang sangat efektif untuk komunikasi politik, mobilisasi pemilih, dan analisis sentimen publik[5]. Dalam konteks pemilihan Presiden, Platform X tidak hanya berfungsi sebagai platform untuk kampanye dan advokasi politik tetapi juga sebagai ruang publik virtual di mana warga negara dapat berpartisipasi dalam diskusi politik secara bebas, mengungkapkan dukungan atau kritik, dan membentuk opini publik[6].

Dengan demikian, analisis sentimen publik yang tercermin dalam interaksi di media sosial menjadi kunci untuk memahami dinamika sosial-politik kontemporer[7]. Namun, tantangan muncul dengan volume data yang sangat besar yang dihasilkan, membuat analisis manual menjadi tidak praktis dan sangat memakan waktu. Ini memperkuat kebutuhan akan peyerngembangan metode otomatis yang

dapat mengklasifikasikan sentimen dari teks yang dihasilkan di media sosial dengan cepat dan akurat. Di era digital saat ini, di mana informasi dapat menyebar dengan sangat cepat dan opini publik dapat dengan mudah dipengaruhi oleh narasi online, kemampuan untuk memahami sentimen publik secara mendalam dan sistematis menjadi sangat penting[9]. Untuk menglola dari volume dari sebuah data text yang berasal dari media social ini tentunya menjadi hal yang rumit dan sulit jika dilakukan secara manual untuk menganalisis hal tersebut. Sehingga perlu adanya penerapan text mining untuk membantu dalam mengolah text tersebut. Text mining adalah sebuah proses ekstraksi terhadap suatu text atau dokumen untuk diolah dengan Bahasa computer untuk mengetahui sentimen atau pendapat[10], khususnya dalam hal ini adalah sentimen terhadap hasil pemilihan presiden 2024.

Untuk mengatasi tantangan tersebut, penelitian ini mengusulkan penggunaan model Algoritma *Support Vector Machine* (SVM) merupakan metode komputasi data dengan dimensi tinggi sehingga tingkat akurasi yang dihasilkan lebih baik[11], sebuah pendekatan revolusioner dalam pemrosesan 81entim alami (NLP) yang memanfaatkan teknik *deep learning* untuk menghasilkan representasi kata yang kaya dan kontekstual. SVM dirancang untuk memahami nuansa dan konteks dalam teks secara lebih efektif daripada model NLP sebelumnya[12]. Implementasi model SVM dalam analisis 81 entiment pasca-pemilihan Presiden memungkinkan identifikasi opini dan emosi dalam teks dengan akurasi yang lebih tinggi, menyajikan peluang untuk analisis yang lebih mendalam dan berwawasan tentang dinamika opini publik[13].

Penelitian sebelumnya telah membahas metode *Support Vector Machine* (SVM) untuk klasifikasi terhadap calon presiden [11]. Penelitian ini fokus pada sentimen terhadap calon Presiden Indonesia berdasarkan tweet. Penelitian ini, sampel diambil dari tiga data tweet yang mencantumkan nama calon Presiden Indonesia 2024. Analisis sentimen dilakukan menggunakan SVM, dan hasilnya menunjukkan nilai akurasi sebesar 75% pada calon 1, 86% pada calon 2, dan 72% pada calon 3. Total dataset yang digunakan dalam penelitian ini adalah sebanyak 1719 tweet. Penelitian ini menunjukkan bahwa SVM dapat menjadi metode yang efektif untuk klasifikasi

sentimen politik di media sosial, meskipun tingkat akurasi bervariasi antar calon.

Penelitian selanjutnya yang dilakukan oleh [14] menganalisis sentimen kata "PILPRES" (Pemilihan Presiden) di Twitter pada tahun 2024, menggunakan dataset 5000 tweet dan memperoleh akurasi 65%. Model ini menunjukkan performa lebih baik dalam mengidentifikasi sampel positif dengan recall 81% dan precision 69%, yang berarti model lebih baik dalam mendeteksi sampel positif meskipun terdapat false positive. F1-score untuk sentimen positif mencapai 74%, menunjukkan keseimbangan antara precision dan recall. Hasil ini mengindikasikan bahwa metode SVM memiliki performa yang relatif baik dalam mengenali dan mengklasifikasikan sentimen terkait PILPRES 2024 di Indonesia.

Kemudian penelitian yang dilakukan oleh [15] ini membandingkan metode Naïve Bayes dengan SVM untuk mengetahui respon terhadap isu penundaan Pemilu 2024 menggunakan 585 data uji. Hasilnya menunjukkan bahwa metode *Support Vector Machine* (SVM) mencapai akurasi sebesar 91,61%, sedangkan Naïve Bayes mencapai akurasi sebesar 98,80%. Dari data tersebut, 429 terprediksi sebagai sentimen negatif dan 155 sebagai sentimen positif. Hal ini menunjukkan bahwa sebagian besar masyarakat tidak mempercayai berita hoax terkait isu penundaan Pemilu 2024.

Penelitian selanjutnya oleh [16] melakukan komparasi klasifikasi Naïve Bayes dengan SVM berdasarkan opini tweet masyarakat tentang Pilpres 2024 yang berasal dari media sosial Twitter. Dataset diambil dari bulan April hingga Mei 2024, dengan total 1606 tweet. Hasil klasifikasi menunjukkan bahwa *Support Vector Machine* (SVM) mendapatkan hasil yang lebih tinggi dengan akurasi 98,43%, precision 97,15%, dan recall 99,71%. Sedangkan algoritma Naïve Bayes memiliki akurasi 96,63%, precision 94,30%, dan recall 98,90%.

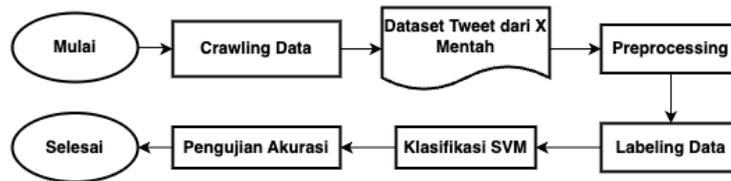
Eksplorasi Analisis Sentimen dalam media sosial khususnya untuk mengetahui perubahan opini publik pasca-pemilu, akan menjadi literatur baru yang bernilai. Meskipun SVM telah terbukti menjadi salah satu metode yang efektif dalam banyak kasus analisis sentimen, penjelasan mengenai pemilihannya dalam konteks opini publik pasca-pemilihan Presiden belum tersedia secara komprehensif. Oleh karena itu, penelitian ini bertujuan untuk mengisi celah pengetahuan ini dengan mengeksplorasi aplikasi SVM dalam analisis sentimen media sosial pasca-pemilihan Presiden secara lebih mendalam. Sehingga, penelitian ini tidak hanya akan memberikan wawasan baru tentang penggunaan SVM dalam konteks ini tetapi juga dapat memberikan pemahaman yang lebih baik tentang keunggulan dan potensi keterbatasan metode ini dalam menganalisis sentimen publik pasca-pemilihan presiden.

Permasalahan utama dalam penelitian ini adalah memahami perubahan opini publik pasca-pemilihan presiden melalui analisis sentimen media sosial menggunakan klasifikasi SVM. Bagaimana sentimen publik berubah sesudah pemilu dan respon masyarakat dengan kata yang sedang tren di media sosial X, serta efektivitas media sosial sebagai platform untuk mengukur sentimen tersebut, merupakan fokus penting. Selain itu, penelitian ini mengevaluasi kinerja SVM. Keakuratan dan keterandalan data media sosial, termasuk penanganan akun-akun bot dan kampanye disinformasi, juga menjadi perhatian utama. Tantangan teknis dalam pemrosesan bahasa alami (NLP) pada teks media sosial yang seringkali informal turut dikaji. Penelitian ini bertujuan memberikan wawasan tentang dinamika opini publik pasca-pemilu dan mengevaluasi efektivitas serta tantangan dalam menggunakan analisis sentimen media sosial sebagai alat pengukur sentimen publik [17].

2. METODE PENELITIAN

2.1 Alur Penelitian

Untuk memudahkan penelitian, diperlukan kerangka atau alur sebagai panduan dalam prosesnya. Ini akan membantu memastikan langkah-langkah yang efisien dalam pengumpulan informasi dan pengolahan data, sehingga hasilnya penelitian ini dan analisis sentimenya menjadi optimal.



Gambar 1. Alur Penelitian

2.2 Pengumpulan Data

Penelitian ini menggunakan metode analisis sentimen untuk mengevaluasi opini publik terhadap Pemilu Presiden 2024, dengan memanfaatkan kumpulan tweet yang dikumpulkan dalam rentang waktu antara tanggal 14 Februari hingga 13 Maret 2024. Sebanyak 3850 tweet telah dikumpulkan melalui pencarian menggunakan kata kunci terkait seperti "Pemilu Damai", "Hak Angket", dan "Pemilu Curang". Proses pengumpulan data dilakukan dengan menggunakan bahasa pemrograman Python dan IDE Google Colab, memungkinkan analisis yang efisien dan cepat. Langkah-langkah analisis sentimen meliputi pra-pemrosesan teks, ekstraksi fitur, dan penerapan model analisis sentimen, dengan tujuan untuk memahami opini publik terhadap pasca Pemilu Presiden 2024 serta potensi dampaknya terhadap stabilitas politik dan sosial.

2.3 Teknik Analisis Data

2.3.1 Preprocessing

Sebelum memasuki proses klasifikasi dan pelabelan kumpulan data tweet harus melalui tahapan *preprocessing* untuk mentransformasi data yang awalnya tidak terstruktur atau terdapat banyak noise menjadi data yang terstruktur[18]. Hal ini bertujuan untuk memudahkan proses analisis selanjutnya. Beberapa metode yang akan diterapkan dalam proses data *preprocessing* meliputi tokenisasi teks, penghapusan tanda baca, normalisasi teks

(seperti konversi huruf kecil semua), penghapusan kata-kata yang tidak relevan (*stopwords*), serta *stemming* atau *lemmatisasi* untuk mengubah kata-kata menjadi bentuk dasar. [20]. Setelah tahap *preprocessing* selesai, data yang dihasilkan akan siap untuk masuk ke dalam proses klasifikasi sentimen[21].

2.3.2 Pelabelan Data

Pelabelan data dilakukan dengan mengklasifikasikan setiap tweet ke dalam tiga kategori sentimen: negatif, netral, dan positif, terkait dengan respon pasca-pemilu Presiden 2024. Proses pelabelan ini dilakukan menggunakan program Python yang membandingkan tiga metode berbeda: NLTK, TextBlob, dan Vader, untuk menilai setiap tweet dan menentukan apakah bersifat negatif, netral, atau positif terhadap konteks pasca Pemilu Presiden 2024. Setelah itu, akan dipilih hasil metode pelabelan yang memiliki akurasi tertinggi terhadap data[22]. Hasil dari pelabelan ini menghasilkan dataset dengan label sentimen yang dapat digunakan untuk analisis lanjutan, sehingga memberikan pemahaman yang lebih baik tentang opini publik terhadap pasca-pemilihan umum Presiden 2024 berdasarkan data tweet yang dikumpulkan.

2.4 Support Vector Machine (SVM)

Pemodelan *Support Vector Machine* (SVM) digunakan untuk menganalisis data dengan cara menemukan batas keputusan optimal antara kelas-kelas yang berbeda. SVM bertujuan untuk memaksimalkan jarak antara titik data dan hyperplane pemisah, sehingga menghasilkan model prediksi yang akurat dan efisien untuk klasifikasi[23]. Untuk memastikan analisis yang optimal, dataset dibagi menjadi dua tipe 80% untuk pelatihan dan 20% untuk pengujian, yang menjadi standar untuk kinerja

klasifikasi[24]. SVM menggunakan pendekatan fungsi kernel untuk melakukan modifikasi matematis, memungkinkannya untuk menemukan hyperplane yang mencapai pemisahan maksimum antara dua kelas, dikenal sebagai hyperplane dengan margin terbesar[25]. Berikut adalah Persamaan dari Hyperlane dari SVM:

$$w \cdot x + b = 0 \quad (1)$$

Dengan w adalah vektor normal terhadap hyperplane. Sedangkan b adalah bias atau pergeseran *hyperplane*.

2.5 Pengukuran Evaluasi SVM

Untuk menilai seberapa efektif klasifikasi yang dihasilkan oleh machine learning, evaluasi diperlukan. Evaluasi ini sering kali dilakukan menggunakan matriks yang disebut *Confusion Matrix*. Matrik ini memberikan gambaran tentang kinerja model dengan mengukur jumlah prediksi yang benar dan salah untuk setiap kelas atau label yang diprediksi.

Tabel 1: Tabel Confusion Matrix

		Kelas Prediksi	
		Positif	Negatif
Observasi	Positif	TP	FN
	Negative	FP	TN

Keterangan :

True Positif (TP) menunjukkan prediksi yang benar untuk label positif, sementara True Negatif (TN) menunjukkan prediksi yang benar untuk label negatif. False Positif (FP) terjadi ketika prediksi label positif ternyata salah, sedangkan False Negatif (FN) terjadi ketika prediksi label negatif ternyata salah. Confusion Matrix memvisualisasikan konsep sebagai dasar evaluasi klasifikasi.

Selain itu, perhitungan akurasi, presisi, recall dan F1-Score juga dilakukan untuk mengetahui hasil dari metode SVM. Untuk melakukan perhitungannya digunakan beberapa rumus matematis sebagai berikut[26]:

$$\text{Akurasi} = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

$$F1 - \text{Score} = \frac{2 \cdot (\text{Recall} \cdot \text{Presisi})}{\text{Recall} + \text{Presisi}} \quad (5)$$

3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk menganalisis sentimen terhadap kumpulan tweet dari platform media sosial X menggunakan tiga kata dari *search engine* yang sedang tren setelah pemilu, yaitu Hak Angket, Pemilu Damai, dan Pemilu Curang. Tiga kata kunci tersebut akan digunakan sebagai acuan dalam membandingkan hasil respon dari masyarakat terhadap apa yang terjadi dan yang menjadi tren di media sosial X. Tujuannya adalah untuk memahami tanggapan masyarakat terhadap isu yang muncul pasca pemilihan presiden Indonesia 2024. Diharapkan penelitian ini dapat memberikan wawasan yang lebih dalam tentang pandangan publik.

3.1 Pengumpulan Data

Pada tahap ini, ditampilkan data yang telah diperoleh pada tahap pengambilan data yang berasal dari scraping dari platform X yang menggunakan search engine untuk mencari data tweet yang dibutuhkan. Data tersebut mencakup data sekunder yaitu yang berisi hasil tweet tentang respon masyarakat terkait hasil pemilihan umum Presiden 2024 di Indonesia. Jumlah data yang dikumpulkan dalam penelitian ini sebanyak 3850 data. Dalam proses pengambilan data memanfaatkan 3 kata kunci search engine yaitu Pemilu Damai memperoleh 850 dataset, Pemilu Curang 1500 Dataset, dan Hak Angket mendapatkan 1500 Dataset yang diambil dari 14 Februari – 13 Maret 2024. Berikut adalah sampel hasil dari pengumpulan dataset dari tweet platform X :

Tabel 2: Dataset mentah

No	Tweet
1	Masyarakat dihimbau agar tidak terprovokasi dengan hasil real count maupun quick count #KPU #Bawaslu #Pemilu2024 #PemiluDamai #IndonesiaMaju https://t.co/oqhudlgoJR
2	@DoankWarto Pemilu kali ini paling sejuk,damai dan penuh kegembiraan.
3	Tunggu hasil resmi KPU. Mari sukseskan Pemilu 2024 dengan Aman dan damai. #pilpres2024 #PemiluDamai2024 #Pemilu2024 https://t.co/gn1hq8aR2D
850	Pemilu Damai Indonesia Maju: Kunci Keberhasilan Demokrasi Berkelanjutan https://t.co/MRXVcREQv9

Proses pengumpulan dataset dari platform media sosial X menggunakan metode scraping dengan Python dan Google Colab dapat dilakukan dengan menggunakan beberapa library seperti menggunakan metode harverst dengan menggunakan token dari platform X sehingga tidak terbatas oleh limitasi.

3.2 Preprocessing

Preprocessing merupakan tahapan awal untuk mempersiapkan dan membersihkan data mentah yang sudah kita ambil agar sesuai dengan kebutuhan analisis lebih lanjut. Langkah-langkah *preprocessing* meliputi normalisasi, penghapusan data yang tidak relevan atau *noise*, serta pengelompokan atau penggabungan data yang serupa[27]. Melalui proses ini, informasi yang bermutu dapat diekstraksi dari dataset yang kompleks, memungkinkan peneliti untuk menerapkan metode analisis yang tepat dengan hasil yang lebih konsisten dan dapat diandalkan. *Preprocessing* juga bertujuan untuk meningkatkan kualitas data, mengurangi bias, dan memfasilitasi interpretasi yang akurat[28]. Adapun beberapa tahapan antara lain :

a. Case folding

Case folding adalah proses mengubah semua huruf dalam teks menjadi huruf kecil. Hal ini dilakukan untuk membuat perbandingan teks menjadi lebih konsisten, terlepas dari apakah huruf-hurufnya awalnya ditulis dalam huruf besar atau kecil.

Tabel 3: Case Folding

Tweet	Case Folding
Masyarakat dihimbau agar tidak terprovokasi dengan hasil real count maupun quick count #KPU #Bawaslu #Pemilu2024 #PemiluDamai #IndonesiaMaju https://t.co/oqhudlgoJR	masyarakat dihimbau agar tidak terprovokasi dengan hasil real count maupun quick count kpu bawaslu pemilu pemiludamai indonesiamaju

b. Cleaning

Cleaning adalah proses menghapus atau memodifikasi data yang tidak relevan atau tidak diinginkan dari sebuah dataset

atau teks. Tujuannya adalah untuk membersihkan data sehingga lebih mudah untuk dianalisis atau diproses.

Tabel 4: Cleaning

Tweet	Cleaning
Masyarakat dihimbau agar tidak terprovokasi dengan hasil real count maupun quick count #KPU #Bawaslu #Pemilu2024 #PemiluDamai #IndonesiaMaju https://t.co/oqhudlgoJR	masyarakat dihimbau agar tidak terprovokasi dengan hasil real count maupun quick count kpu bawaslu pemilu pemiludamai indonesiamaju

c. Tokenize

Tokenize adalah proses membagi teks menjadi unit-unit yang lebih kecil, seperti kata, frasa, atau karakter terpisah, untuk memudahkan analisis atau pemrosesan data.

Tabel 5: Tokenize

Tweet	Tokenize
masyarakat dihimbau agar tidak terprovokasi dengan hasil real count maupun quick count kpu bawaslu pemilu pemiludamai indonesiamaju	[masyarakat, dihimbau, agar, tidak, terprovokasi, dengan, hasil, real, count, maupun, quick, count, kpu, bawaslu, pemilu, pemiludamai, indonesiamaju,]

d. Penghapusan Stopword

Penghapusan *Stopword* adalah proses menghapus kata-kata umum yang tidak memberikan nilai tambah dalam analisis teks.

Tabel 6: Stopwords

Tweet	Stopwords
masyarakat dihimbau agar tidak terprovokasi dengan hasil real count maupun quick count kpu bawaslu pemilu	[masyarakat, dihimbau, terprovokasi, hasil, real, count, quick, count, kpu, bawaslu, pemilu,

pemiludamai indonesiamaju	pemiludamai, indonesiamaju,]
------------------------------	----------------------------------

e. *Lemmatized*

Lemmatized adalah proses mengubah kata-kata dalam teks menjadi bentuk dasarnya (lemmata) untuk mempermudah analisis. Misalnya, kata-kata seperti "berlari", "berlari", dan "lari" akan diubah menjadi bentuk dasar "lari". Tujuan utamanya adalah untuk mengurangi variasi kata yang serupa sehingga analisis teks menjadi lebih konsisten dan akurat.

Tabel 7: *Lemmatized*

<i>Tweet</i>	<i>Lemmatized</i>
masyarakat dihimbau agar tidak terprovokasi dengan hasil real count maupun quick count kpu bawaslu pemilu pemiludamai indonesiamaju	[masyarakat, dihimbau, provokasi, hasil, real, count, quick, count, kpu, bawaslu, milu, pemiludamai, indonesiamaju,]

f. Normalisasi

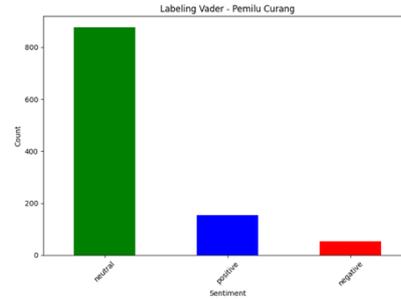
Normalisasi adalah proses membuat data seragam atau standar untuk mempermudah analisis atau pemrosesan lebih lanjut.

Tabel 8: Normalisasi

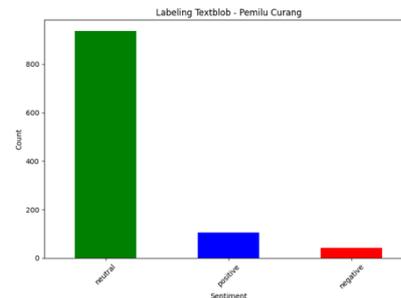
<i>Tweet</i>	<i>Normalisasi</i>
masyarakat dihimbau agar tidak terprovokasi dengan hasil real count maupun quick count kpu bawaslu pemilu pemiludamai indonesiamaju	masyarakat dihimbau provokasi hasil real count quick count kpu bawaslu milu pemiludamai indonesiamaju

3.3 Pelabelan data

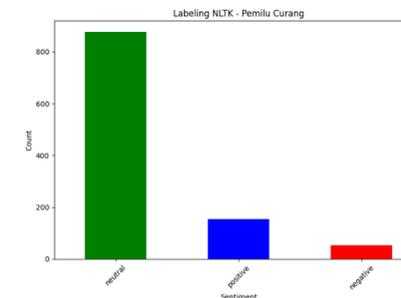
Setelah dilakukan preprocessing, dataset dilabeli dengan klasifikasi menjadi tiga sentimen yaitu positif, netral, dan negatif. Proses pelabelan ini menggunakan fungsi dari NLTK, TextBlob, dan Vader untuk mengetahui akurasi hasilnya. Dari hasil evaluasi, diketahui bahwa fungsi NLTK dan Vader memiliki akurasi yang hampir mirip. Oleh karena itu, diputuskan untuk menggunakan fungsi NLTK secara otomatis dalam pelabelan dataset ini.



Gambar 2. Labeling Vader



Gambar 3. Labeling Textblob



Gambar 4. Labeling NLTK

3.4 Visualisasi Word Cloud

Tahapan ini berfungsi memvisualisasikan dataset kata yang sering muncul menjadi sebuah grafis untuk mengetahui identifikasi data sehingga lebih mudah dilakukan analisis. Dari tiga search yang kita lakukan analisis sesuai dengan dataset yaitu "Pemilu Damai", "Pemilu Curang", dan "Hak Angket". Proses word cloud menggunakan library matplotlib dalam bahasa pemrograman Python.

Model klasifikasi dari Dataset “Hak Angket” menunjukkan bahwa model memiliki tingkat akurasi sebesar 97% Dari 259 tweet yang diklasifikasikan. Mayoritas (217 tweet) dikategorikan sebagai positif dengan tingkat recall dan presisi sebesar 98%, menunjukkan kemampuan model dalam mengidentifikasi tweet yang menyuarakan pendapat positif terkait Hak Angket dengan sangat baik. Namun, model memiliki tingkat recall yang rendah untuk kategori negatif (33%) yang menunjukkan kesulitan dalam mengidentifikasi tweet-tweet yang mengungkapkan pendapat negatif terkait Hak Angket. Sementara itu, kategori netral memiliki tingkat recall dan presisi yang cukup tinggi (95% dan 88% berturut-turut), menunjukkan bahwa model cukup baik dalam mengenali tweet yang netral terkait Hak Angket.

	precision	recall	f1-score	support
negative	0.00	0.00	0.00	2
neutral	0.97	1.00	0.98	64
positive	1.00	1.00	1.00	9
accuracy			0.97	75
macro avg	0.66	0.67	0.66	75
weighted avg	0.95	0.97	0.96	75

Gambar 11. Klasifikasi Data Pemilu Damai

Hasil klasifikasi dari dataset “Pemilu Damai” menunjukkan akurasi sebesar 97% dalam mengklasifikasikan tweet terkait pemilu damai menjadi negatif, netral, atau positif. Total tweet yang dilakukan testing adalah 67 dikategorikan sebagai netral dengan tingkat presisi dan recall 97% dan 100% berturut-turut. Namun, kategori negatif memiliki nilai recall 0%, menunjukkan kesulitan model dalam mengidentifikasi tweet-tweet yang mengungkapkan pendapat negatif terkait pemilu damai. Kategori positif memiliki presisi dan recall 100%.

Selain itu, dilakukan berikut adalah hasil dari sentimen analisis menggunakan metode SVM :

Tabel 9: Implementasi Labeling SVM

Text	Sentimen
nah yuk kita buktikan di angket biar terbuka siapa penjagat suara rakyat	Positif
biasa aja itu tandanya pemilu sudah mamasuki babak baru ya angket bukan soal capres lagi	Netral
culas curang kotor dan brutal kejahatan pilpres	Negatif

3.7 Akurasi

Akurasi adalah metrik yang mengukur seberapa tepat model klasifikasi dalam mengklasifikasikan data. Dalam konteks yang diberikan, terdapat tiga model yang dinilai:

- Model untuk Klasifikasi Hak Angket memiliki akurasi sebesar 0.965. Ini berarti bahwa sekitar 96.5% dari seluruh data yang digunakan untuk menguji model telah diklasifikasikan dengan benar sebagai positif, netral, atau negatif terkait dengan Hak Angket.

```
accuracy_score(y_test, prediction)
0.9652509652509652
```

- Model untuk Klasifikasi Pemilu Damai memiliki akurasi sebesar 0.973. Ini menunjukkan bahwa sekitar 97.3% dari seluruh data yang digunakan untuk menguji model telah diklasifikasikan dengan benar sebagai positif, netral, atau negatif terkait dengan Pemilu Damai.

```
accuracy_score(y_test, prediction)
0.9733333333333334
```

- Model untuk Klasifikasi Pemilu Curang memiliki akurasi sebesar 0.940. Ini berarti sekitar 94.0% dari seluruh data yang digunakan untuk menguji model telah diklasifikasikan dengan benar sebagai positif, netral, atau negatif terkait dengan Pemilu Curang.

```
accuracy_score(y_test, prediction)
0.9400921658986175
```

4. KESIMPULAN

Hasil klasifikasi sentimen pasca-pemilihan presiden yang dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dengan data yang dikumpulkan dari platform media sosial X adalah sebagai berikut. Dari total dataset sebanyak 3850, ditemukan bahwa model untuk klasifikasi "Pemilu Damai" memiliki kinerja terbaik dengan akurasi 97%, diikuti oleh model untuk "Hak Angket" dengan akurasi 96,5%, dan model untuk "Pemilu Curang" dengan akurasi 94,0%. Namun,

penting untuk dicatat bahwa hasil akurasi dari ketiga kelas tersebut juga dipengaruhi oleh jumlah dataset yang ada, di mana semakin sedikit data yang tersedia, semakin tinggi pula hasil akurasinya. Hal ini menunjukkan perlunya perhatian terhadap keseimbangan jumlah data antar kelas dalam proses analisis sentimen, serta menekankan pentingnya mempertimbangkan faktor ini dalam evaluasi kinerja model klasifikasi sentimen.

Berdasarkan hasil analisis sentimen terhadap pasca pemilihan presiden 2024, saran yang dapat diberikan adalah fokus pada peningkatan kinerja model untuk menghasilkan klasifikasi yang optimal. Perbaikan dapat dilakukan dengan memperluas dataset latihan, melakukan pemrosesan teks yang lebih cermat, dan menyesuaikan parameter model. Selain itu, penggunaan teknik pemodelan yang lebih canggih seperti penggunaan model berbasis deep learning atau ensemble learning juga dapat dipertimbangkan. Pentingnya evaluasi terus-menerus terhadap kinerja model juga harus ditekankan.

DAFTAR PUSTAKA

- [1] Z.- Arman, "ANALISIS TERHADAP PENGUJIAN UNDANG-UNDANG NOMOR 42 TAHUN 2008 TENTANG PEMILIHAN UMUM PRESIDEN DAN WAKIL PRESIDEN," *Jurnal Cahaya Keadilan*, vol. 7, no. 1, p. 264, Apr. 2019, doi: 10.33884/jck.v7i1.1199.
- [2] D. P. Ray, F. N. Hasan, and A. R. Dzikrillah, "KLIK: Kajian Ilmiah Informatika dan Komputer Analisis Sentimen Terhadap KPU 2024 Berdasarkan Tweet Media Sosial Twitter Menggunakan Algoritma Naïve Bayes," *Media Online*, vol. 4, no. 4, pp. 2235–2243, 2024, doi: 10.30865/klik.v4i4.1587.
- [3] W. Winarto, I. Alwiah Musdar, P. Studi Teknik Informatika, S. Kharisma Makassar, P. Studi Sistem Informasi, and U. Alauddin Makassar, "ANALISIS SENTIMEN BAKAL CALON PRESIDEN 2024 MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE PADA TWITTER", [Online]. Available: <https://jurnal.kharisma.ac.id/kharismatech>
- [4] N. Prisca and A. 1✉, "Transformasi Media Sosial Dalam Kompetensi Komunikasi Politik," *INNOVATIVE: Journal Of Social Science Research*, vol. 3, pp. 5101–5109, 2023.
- [5] P. Arsi and R. Waluyo, "ANALISIS SENTIMEN WACANA PEMINDAHAN IBU KOTA INDONESIA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)," vol. 8, no. 1, pp. 147–156, 2021, doi: 10.25126/jtiik.202183944.
- [6] S. A. Rachmini, I. Rusman, D. M. Sari, A. Irianti, and S. Saparang, "Analisis Sentimen Publik Terkait Bakal Calon Presiden 2024 Menggunakan Metode Support Vector Machine," 2023.
- [7] A. Arsyad, R. Ghulam Dzaljad, M. Nurmiarani, and S. Rantona, "Media Sosial sebagai Agen Transformasi Politik: Analisis Pengaruh terhadap Proses Komunikasi Politik," *Jurnal Pendidikan Dan Ilmu Sosial*, vol. 2, no. 2, pp. 240–251, 2024, doi: 10.54066/jupendis.v2i2.1593.
- [8] I. P. Dedy, W. Darmawan, G. Aditra Pradnyana, I. Bagus, and N. Pascima, "Optimasi Parameter Support Vector Machine Dengan Algoritma Genetika Untuk Analisis Sentimen Pada Media Sosial Instagram", [Online]. Available: <https://doi.org/10.31598>
- [9] R. N. Mauliza and Y. R. Sipayung, "Penerapan Text Mining Dalam Menganalisis Pendapat Masyarakat Terhadap Pemilu 2024 Pada Media Sosial X Menggunakan Metode Naive Bayes," *Technomedia Journal (TMJ)*, vol. 9, no. 1, p. 9, 2024, doi: 10.33050/tmj.v9i1.2212.
- [10] A. Mustofa and R. Novita, "Klasifikasi Sentimen Masyarakat Terhadap Pemberlakuan Pembatasan Kegiatan Masyarakat Menggunakan Text Mining Pada Twitter," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 1, Jun. 2022, doi: 10.47065/bits.v4i1.1628.
- [11] A. Handayani and I. Zufria, "Analisis Sentimen Terhadap Bakal Capres RI 2024 di Twitter Menggunakan Algoritma SVM," *Journal of Information System Research (JOSH)*, vol. 5, no. 1, pp. 53–63, Oct. 2023, doi: 10.47065/josh.v5i1.4379.
- [12] P. Kinerja Naïve Bayes, A. Supian, B. Tri Revaldo, N. Marhadi, and L. Efrizoni,

- “Acuan Supian Perbandingan Kinerja Naïve Bayes dan SVM pada Analisis Sentimen Twitter Ibukota Nusantara.” [Online]. Available: <https://github.com/syenirasheila/Sentiment-Analysis-IKN->
- [13] D. Darwis, E. Shintya Pratiwi, A. Ferico, and O. Pasaribu, “PENERAPAN ALGORITMA SVM UNTUK ANALISIS SENTIMEN PADA DATA TWITTER KOMISI PEMBERANTASAN KORUPSI REPUBLIK INDONESIA,” 2020.
- [14] N. Alexander, R. Bria, and A. Witanti, “ANALISIS SENTIMEN MASYARAKAT INDONESIA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE TENTANG PILPRES 2024,” 2023.
- [15] P. Guru *et al.*, “Analisis Sentimen Masyarakat Terhadap Isu Penundaan Pemilu 2024 Pada Twitter Dengan Metode Naïve Bayes Dan Support Vector Machine,” *Jurnal Sains dan Teknologi*, vol. 5, no. 3, p. 2024, doi: 10.55338/saintek.v5i1.2789.
- [16] T. Rosyida, H. P. Putro, and H. Wahyono, “ANALISIS SENTIMEN TERHADAP PILPRES 2024 BERDASARKAN OPINI DARI TWITTER MENGGUNAKAN NAÏVE BAYES DAN SVM”, [Online]. Available: www.apjii.or.id
- [17] O. Manullang and C. Prianto, “Analisis Sentimen dalam Memprediksi Hasil Pemilu Presiden dan Wakil Presiden : Systematic Literature Review”, [Online]. Available: <https://ejurnalunsam.id/index.php/jicom/>
- [18] I Putu Dedy Wira Darmawan, Gede Aditra Pradnyana, and Ida Bagus Nyoman Pascima, “Optimasi Parameter Support Vector Machine Dengan Algoritma Genetika Untuk Analisis Sentimen Pada Media Sosial Instagram”, [Online]. Available: <https://doi.org/10.31598>
- [19] T. Meisya, P. Aulia, N. Arifin, and R. Mayasari, “PERBANDINGAN KERNEL SUPPORT VECTOR MACHINE (SVM) DALAM PENERAPAN ANALISIS SENTIMEN VAKSINISASI COVID-19”, [Online]. Available: <https://doi.org/10.31598>
- [20] Shahmirul Hafizullah Imanuddin, Kusworo Adi, and Rahmat Gernowo, “Sentiment Analysis on Satushehat Application Using Support Vector Machine Method,” *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 5, no. 3, pp. 143–149, Jul. 2023, doi: 10.35882/jeemi.v5i3.304.
- [21] E. Suryati, A. Ari Aldino, N. Penulis Korespondensi, and E. Suryati Submitted, “Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM),” vol. 4, no. 1, pp. 96–106, 2023, doi: 10.33365/jtsi.v4i1.2445.
- [22] P. Kumala Sari and R. Randy Suryono, “KOMPARASI ALGORITMA SUPPORT VECTOR MACHINE DAN RANDOM FOREST UNTUK ANALISIS SENTIMEN METAVERSE,” 2024.
- [23] C. K. Wang, “Sentiment Analysis Using Support Vector Machines, Neural Networks, and Random Forests,” 2023, pp. 23–34. doi: 10.2991/978-94-6463-300-9_4.
- [24] A. Rolangon *et al.*, “Perbandingan Algoritma LSTM Untuk Analisis Sentimen Pengguna Twitter Terhadap Layanan Rumah Sakit Saat Pandemi Covid-19 The Comparison of LSTM Algorithms for Twitter User Sentiment Analysis on Hospital Services During the Covid-19 Pandemic.”
- [25] W. Ningsih, B. Alfiana, R. Rahmaddeni, and D. Wulandari, “Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Twitter pada Penggunaan Mobil Listrik di Indonesia,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 556–562, Feb. 2024, doi: 10.57152/malcom.v4i2.1253.
- [26] Ni Wayan Sumartini Saraswati, I Putu Krisna Suarendra Putra, I Dewa Made Krishna Muku, and Gede Dana Pramitha, “Support Vector Machine For Hoax Detection”, [Online]. Available: <https://doi.org/10.31598>
- [27] H. Mustofa and A. A. Mahfudh, “Klasifikasi Berita Hoax Dengan Menggunakan Metode Naïve Bayes,” *Walisongo Journal of Information Technology*, vol. 1, no. 1, p. 1, Nov. 2019, doi: 10.21580/wjit.2019.1.1.3915.

- [28] C. Huda and M. Betty Yel, "Analisa Sentimen Tentang Ibu Kota Nusantara (IKN) Dengan Menggunakan Algoritma K-Nearest Neighbors (KNN) dan Naïve Bayes," *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI V*, vol. 7, no. 1, pp. 126–130, 2024.