ANALISIS SENTIMEN MASYARAKAT TERHADAP DANANTARA DI PLATFORM X DENGAN METODE SVM

Muhammad Lucky Hermanto, Fathoni, Onky Ardhillah, Ali Ibrahim
Fakultas Ilmu Komputer, Universitas Sriwijaya, Jl. Raya Palembang - Prabumulih KM. 32, Indralaya Indah,Kec. Indralaya, Kabupaten Ogan Ilir, Sumatera Selatan 30862
09031382227180@student.unsri.ac.id

ABSTRAK

Danantara merupakan lembaga pengelola investasi baru yang diluncurkan oleh pemerintah Indonesia dan telah menjadi perbincangan hangat di media sosial, khususnya platform X (sebelumnya Twitter). Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap Danantara dengan menggunakan metode Support Vector Machine (SVM). Sebanyak 1.175 tweet yang mengandung kata kunci "Danantara" dikumpulkan dari Januari hingga April 2025. Setelah melalui tahapan preprocessing seperti pembersihan teks, tokenisasi, penghapusan stopword, stemming, dan ekstraksi fitur menggunakan TF-IDF, dilakukan pelabelan otomatis dan klasifikasi sentimen menjadi positif dan negatif. Hasil evaluasi menunjukkan bahwa model SVM mampu mengklasifikasikan sentimen dengan tingkat akurasi sebesar 79,57%, precision 79,54%, recall 69,62%, dan F1-score 71,63%. Visualisasi dengan word cloud mengungkapkan kata-kata dominan seperti "koruptor", "asing", dan "danantara", yang menunjukkan fokus perhatian publik. Penelitian ini memberikan wawasan penting bagi pengambil kebijakan dalam memahami opini publik serta menyusun strategi komunikasi yang lebih efektif terhadap program Danantara.

Kata kunci : *Danantara, analisis sentimen, Support Vector Machine, media sosial, Twitter*

1. PENDAHULUAN

Lembaga Pengelola Investasi Daya Anagata Nusantara, yang biasa disebut Danantara Indonesia, memainkan peran penting dalam mengonsolidasikan investasi pemerintah dengan tujuan utama untuk merangsang dan meningkatkan pertumbuhan ekonomi nasional. Didirikan di bawah naungan Presiden Prabowo Subianto, nomenklatur lembaga ini memiliki makna yang signifikan: "Daya," yang berarti energi, mewujudkan pendekatan dinamis lembaga ini terhadap investasi; "Anagata," yang mewakili masa depan, mencerminkan komitmennya untuk mendorong pembangunan berkelanjutan dan inisiatif yang berpikiran maju; dan "Nusantara," istilah yang kaya akan warisan budaya, menggarisbawahi dedikasi lembaga ini untuk menyatukan kepulauan Indonesia yang beragam. Melalui manajemen investasi yang strategis, Danantara Indonesia bertujuan untuk menciptakan lanskap ekonomi yang kuat yang menguntungkan semua lapisan masyarakat [1]. Danantara dibentuk dengan tujuan mengelola dan mengembangkan aset negara guna meningkatkan kinerja serta pengembalian investasi pemerintah, mengikuti model serupa dengan Temasek di Singapura. Pada tahap awal, Danantara akan mengelola tujuh Badan Usaha Milik Negara (BUMN) utama, yaitu Bank Mandiri, Bank Rakyat Indonesia (BRI), PLN, Pertamina, BNI, Telkom Indonesia, dan MIND ID. Total aset yang dikelola oleh badan ini diperkirakan mencapai sekitar Rp 14.678 triliun atau setara dengan US$ 900 miliar [2]. Meskipun diharapkan dapat mendukung pembangunan nasional, keberadaan Danantara juga menimbulkan kekhawatiran terkait risiko mismanajemen dan potensi korupsi. Keberhasilannya akan sangat bergantung pada penerapan standar tata kelola yang baik serta independensi dalam pengelolaannya.

Media sosial, khususnya Twitter, telah menjadi platform populer bagi masyarakat untuk menyampaikan pendapat dan berbagi informasi tentang berbagai topik. Pengguna Twitter dapat dengan bebas mengekspresikan emosi, ide, dan reaksinya. Dengan menggunakan data dari Twitter, analisis sentimen dapat dilakukan untuk mengklasifikasikan opini publik menjadi dua kategori, yaitu positif dan negatif [3]. Analisis sentimen digunakan untuk mengetahui sudut pandang atau pola seseorang mengenai suatu isu atau item tertentu (Zuriel and Fahrurozi, 2021). Analisis sentimen ulasan pengguna sangat membantu karena menyoroti ekspektasi dan pengalaman pengguna, yang dapat berubah dengan cepat [4]. Analisis sentimen berfungsi sebagai alat yang berharga untuk menilai opini publik secara sistematis mengenai Program Danantara milik pemerintah. Dengan mengevaluasi sentimen yang diungkapkan di berbagai forum publik, analisis ini dapat menjelaskan manfaat yang dirasakan dari program tersebut dan dampak sosialnya yang lebih luas.

2. TINJAUAN PUSTAKA

Untuk membantu penelitian ini, kita perlu melihat penelitian dan informasi lain yang membahas topik yang sama. Ini akan memastikan penelitian kita dibangun di atas ide-ide yang baik dan terorganisasi yang benar-benar terkait dengan pertanyaan yang ingin kita jawab.

2.1. Penelitian Terdahulu

Penelitian sebelumnya telah banyak mengeksplorasi penerapan analisis sentimen menggunakan metode pembelajaran mesin, terutama Support Vector Machine (SVM) dalam berbagai konteks sosial dan ekonomi. Penelitian oleh Ananda dan Suryono tahun 2024 membandingkan performa algoritma SVM dan Naïve Bayes dalam analisis sentimen publik terhadap pengungsi Rohingya di Indonesia. Hasilnya menunjukkan bahwa algoritma SVM memberikan performa prediksi yang lebih baik dengan tingkat akurasi 76%, dibandingkan dengan Naïve Bayes yang hanya mencapai 70%. Hasil ini mendukung anggapan bahwa SVM lebih unggul dalam klasifikasi sentimen pada data teks sosial yang kompleks [5].

Sementara itu, Handayani & Zufria tahun 2023 dalam penelitiannya mengenai analisis sentimen terhadap bakal calon presiden Indonesia tahun 2024 di Twitter, menggunakan algoritma SVM dengan tingkat akurasi mencapai 78,3%. Penelitian ini memanfaatkan 1719 tweet dan menunjukkan bahwa algoritma SVM mampu menangani data sosial media dalam jumlah besar serta menghasilkan klasifikasi sentimen yang cukup akurat [6].

Studi lainnya, Arsi & Waluyo tahun 2021 menerapkan SVM untuk menganalisis sentimen publik terhadap isu pemindahan ibu kota negara. Dengan jumlah data sebanyak 1.236 tweet, penelitian ini berhasil memperoleh tingkat akurasi sebesar 96,68% serta nilai AUC sebesar 0,979. Hasil ini menunjukkan bahwa SVM sangat efektif dalam mengekstraksi informasi sentimen dari teks tidak terstruktur di media social [7].

Studi ini berupaya melakukan analisis komprehensif terhadap sentimen publik seputar Program Danantara dengan memanfaatkan teknik pembelajaran mesin tingkat lanjut, khususnya menggunakan algoritma Support Vector Machine (SVM). Penelitian ini bertujuan untuk mengevaluasi kinerja SVM dalam mengklasifikasikan sentimen publik berdasarkan beberapa metrik penting, termasuk akurasi, recall, dan F1-Score. Indikator-indikator ini digunakan untuk menilai efektivitas SVM dalam menangkap opini dan sentimen publik yang bernuansa. Pemahaman menyeluruh tentang persepsi publik sangat penting, karena tidak hanya mempengaruhi keterlibatan pemangku kepentingan tetapi juga memainkan peran penting dalam membentuk keputusan kebijakan terkait Program Danantara. Melalui pendekatan ini, studi ini bertujuan memberikan wawasan yang dapat ditindaklanjuti bagi para pembuat kebijakan dan administrator program, sehingga pada akhirnya dapat meningkatkan efektivitas dan penerimaan inisiatif Danantara. Temuan ini diharapkan dapat memberikan kontribusi yang berarti terhadap wacana yang lebih luas mengenai analisis sentimen publik dan penerapannya dalam evaluasi program serta perencanaan strategis.

2.2. Analisis Sentimen

Analisis sentimen adalah alat yang ampuh yang mengkategorikan sentimen yang diungkapkan dalam teks, seperti ulasan aplikasi, ke dalam klasifikasi positif atau negatif. Proses ini memungkinkan pengembang dan bisnis untuk mendapatkan wawasan berharga tentang opini dan preferensi pengguna, yang pada akhirnya meningkatkan penawaran produk dan kepuasan pelanggan [4].

2.3. Text Mining

Penambangan teks adalah proses analisis canggih yang berfokus pada penggalian wawasan yang bermakna dari kumpulan data teks tak terstruktur yang luas. Teknik ini telah menjadi semakin penting dalam berbagai aplikasi, seperti analisis sentimen dan klasifikasi teks, yang memungkinkan organisasi memperoleh intelijen yang dapat ditindaklanjuti dengan cepat. Dengan menganalisis informasi dari berbagai sumber, termasuk platform media sosial, penambangan teks membekali para pengambil keputusan dengan kemampuan untuk segera menanggapi tren yang muncul dan sentimen publik. Akibatnya, hal ini meningkatkan perencanaan strategis dan efisiensi operasional dalam lanskap informasi yang berkembang pesat [8].

2.4. Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma pembelajaran mesin yang kuat dan serbaguna yang unggul dalam tugas klasifikasi dan regresi. Dengan mengidentifikasi hyperplane yang optimal, SVM memisahkan titik data menjadi dua kelas yang berbeda sekaligus memaksimalkan margin di antara keduanya. Karakteristik ini membuatnya sangat efektif dalam ruang berdimensi tinggi di mana algoritma tradisional mungkin mengalami kesulitan. SVM banyak digunakan dalam berbagai aplikasi, termasuk analisis sentimen, pengenalan gambar, dan bioinformatika. Kemampuannya untuk menangani kumpulan data yang kompleks dan memberikan prediksi yang kuat telah berkontribusi pada popularitasnya di kalangan ilmuwan data dan peneliti, menjadikannya sebagai alat dasar dalam perangkat pembelajaran mesin [4].

2.5. Confusion Matrix

Confusion matrix adalah sebuah tabel yang digunakan untuk menilai kinerja model klasifikasi dengan membandingkan hasil prediksi model dengan nilai aktual pada data uji. Melalui perhitungan confusion matrix, terdapat beberapa metrik evaluasi seperti True Positive (TP) yang menunjukkan jumlah sampel kelas positif yang terklasifikasi dengan benar, dan True Negative (TN) yang menunjukkan jumlah sampel kelas negatif yang juga terklasifikasi dengan benar. Di sisi lain, False Positive (FP) merujuk pada jumlah sampel kelas negatif yang keliru diklasifikasikan sebagai kelas positif, sementara False Negative (FN) adalah jumlah sampel kelas positif yang

salah diklasifikasikan sebagai kelas negatif. Analisis menggunakan confusion matrix ini memungkinkan penghitungan berbagai metrik evaluasi yang digunakan untuk menilai akurasi model klasifikasi [9]. Rumus confusion matrix ditampilkan pada persamaan (1) sampai (4).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (4)$$

3. METODE PENELITIAN

3.1. Pengambilan Data

Data yang digunakan dalam penelitian ini diperoleh dari platform X (sebelumnya dikenal sebagai Twitter), mengingat popularitas platform tersebut dalam menyuarakan opini publik. Pengumpulan data dilakukan dengan menggunakan kata kunci "Danantara". Data diambil pada periode tertentu dari Januari hingga April 2025 untuk memastikan data yang dikumpulkan mencakup berbagai variasi opini masyarakat. yang terkumpul hingga 1175 dataset hasil dari pengambilan data.

Tabel 1. Sampel Data Set

id_str	username	fulltext
620040746	RmDoang	Potensi Danantara untuk dorong ekonomi nasional sangat besar. https://t.co/xDw72ioo5s
809082996	dandyez	@abu_waras simbiosis keknya si kakek mo taruh duit hasil korupnya disini juga lumayan kan dilaundry di danantara
1677082967162900481	natasha_saja	@Mdy_Asmara1701 Demo di tunggangi Asing kata Prabowo. Sementara duit negara yg ada di Danantara dia kasih asing untuk mengelola. Mantan koruptor pula. Ndasmu Asing !

Tabel 1 menyajikan kumpulan data komprehensif yang berkaitan dengan analisis sentimen, yang mencakup elemen-elemen penting seperti ID tweet, nama pengguna, dan teks lengkap setiap tweet. Informasi ini berfungsi untuk menjelaskan sentimen publik mengenai program Danantara, sehingga memudahkan pemahaman yang lebih mendalam tentang persepsi dan opini yang diungkapkan oleh pengguna di platform media sosial mengenai inisiatif ini.

3.2. Pembersihan Data

Pembersihan data mengacu pada proses sistematis untuk menghilangkan atau memodifikasi elemen yang dianggap sebagai gangguan dalam data tekstual, sehingga hanya mempertahankan karakter alfabet yang relevan dengan analisis. Proses ini melibatkan penghapusan komponen yang tidak penting, termasuk tanda baca, angka numerik, simbol, dan karakter alfabet yang tidak relevan [10]. Proses pembersihan data cermat yang diterapkan dalam studi ini memastikan keakuratan dan keandalan temuan RT (retweet), mention (@username), url (http://situs.com), hastag dan simbol-simbol yang tidak diperlukan.

Tabel 2. Data Cleaning

Tweet	Cleaning Data
Potensi Danantara untuk dorong ekonomi nasional sangat besar. https://t.co/xDw72ioo5s	Potensi Danantara untuk dorong ekonomi nasional sangat besar
@abu_waras simbiosis keknya si kakek mo taruh duit hasil korupnya disini juga lumayan kan dilaundry di danantara	simbiosis keknya si kakek mo taruh duit hasil korupnya disini juga lumayan kan dilaundry di danantara
@Mdy_Asmara1701 Demo di tunggangi Asing kata Prabowo. Sementara duit negara yg ada di Danantara dia kasih asing untuk mengelola. Mantan koruptor pula. Ndasmu Asing !	Demo di tunggangi Asing kata Prabowo Sementara duit negara yg ada di Danantara dia kasih asing untuk mengelola Mantan koruptor pula Ndasmu Asing

Tabel 2 menyajikan kumpulan data yang telah disempurnakan, diproses dengan cermat untuk menghilangkan komponen yang tidak relevan seperti URL dan tanda baca. Prosedur pembersihan menyeluruh ini secara signifikan meningkatkan kualitas teks, sehingga lebih sesuai untuk analisis sentimen. Dengan menghilangkan simbol yang mengganggu dan huruf kapital yang berlebihan, data menjadi lebih jelas dan lebih koheren, sehingga memudahkan pemeriksaan sentimen yang lebih akurat. Langkah-langkah praproses tersebut sangat penting untuk memastikan bahwa analisis didasarkan pada konten yang relevan, yang pada akhirnya menghasilkan hasil yang lebih andal dan mendalam.

3.3. Pelabelan Data

Pelabelan data dilakukan secara otomatis memanfaatkan pustaka Deberta Python, yang secara

efisien mengklasifikasikan informasi tekstual dengan memanfaatkan teknik pemrosesan bahasa alami yang canggih untuk meningkatkan organisasi dan analisis data.

Tabel 3. Hasil *Labeling*

Tweet	Label
Potensi Danantara untuk dorong ekonomi nasional sangat besar	Positif
simbiosis keknya si kakek mo taruh duit hasil korupnya disini juga lumayan kan dilaundry di danantara	Negatif
Demo di tunggangi Asing kata Prabowo Sementara duit negara yg ada di Danantara dia kasih asing untuk mengelola Mantan koruptor pula Ndasmu Asing	Negatif

Tabel 3 memberikan gambaran menyeluruh tentang pelabelan sentimen yang terkait dengan tweet yang dianalisis. Tweet pertama, yang menyoroti potensi ekonomi Danantara, dikategorikan sebagai Positif, yang mencerminkan pandangan optimis tentang prospek ekonomi kawasan tersebut. Sebaliknya, tweet kedua dan ketiga, yang membahas masalah korupsi dan implikasi pengelolaan dana negara oleh pihak asing, keduanya diberi label Negatif. Perbedaan ini menggarisbawahi kompleksitas sentimen publik, khususnya yang berkaitan dengan isu ekonomi dan tata kelola yang sangat menyentuh masyarakat.

3.4. Preprocessing Data

Tahap praproses meliputi penghapusan sistematis elemen-elemen yang tidak diperlukan yang dapat menghalangi analisis [11]. Studi ini dengan cermat menggambarkan beberapa tahap penting yang terlibat dalam proses penelitian;

3.5. Case Folding

Case Folding, proses mengubah semua kata berhuruf kapital menjadi huruf kecil dapat meningkatkan keterbacaan secara signifikan.

Tabel 4. Hasil *Case Folding*

Twitter	Case Folding
Potensi Danantara untuk dorong ekonomi nasional sangat besar	potensi danantara untuk dorong ekonomi nasional sangat besar
Demo di tunggangi Asing kata Prabowo Sementara duit negara yg ada di Danantara dia kasih asing untuk mengelola Mantan koruptor pula Ndasmu Asing	demo di tunggangi asing kata prabowo sementara duit negara yg ada di danantara dia kasih asing untuk mengelola mantan koruptor pula ndasmu asing

Tabel 4 mengilustrasikan proses pelipatan huruf, langkah penting dalam praproses data tekstual untuk analisis. Teknik ini melibatkan konversi semua huruf kapital ke huruf kecil, dengan demikian memastikan keseragaman dan konsistensi di seluruh kumpulan data. Misalnya, nama "Danantara" diubah menjadi "danantara," sementara "Asing" diubah menjadi "asing." Standardisasi semacam itu penting, karena meminimalkan perbedaan yang mungkin timbul dari variasi dalam huruf besar-kecil, yang pada akhirnya meningkatkan akurasi dan keandalan metode analisis teks berikutnya.

3.6. Tokenizing

Tokenisasi adalah proses mendasar dalam analisis teks yang melibatkan penguraian teks tertentu menjadi unit-unit yang lebih kecil yang dikenal sebagai token. Segmentasi ini memungkinkan pemeriksaan data linguistik secara sistematis dengan mengisolasi kalimat-kalimat ke dalam komponen-komponen pentingnya. Pendekatan terstruktur seperti itu sangat penting untuk tugas-tugas selanjutnya, termasuk stemming, yang melibatkan pengurangan kata-kata ke bentuk akarnya untuk meningkatkan efisiensi pemrosesan teks [12].

Tabel 5. Hasil *Tokenizing*

Twitter	Tokenizing
potensi danantara untuk dorong ekonomi nasional sangat besar	['potensi', 'danantara', 'untuk', 'dorong', 'ekonomi', 'nasional', 'sangat', 'besar']
demo di tunggangi asing kata prabowo sementara duit negara yg ada di danantara dia kasih asing untuk mengelola mantan koruptor pula ndasmu asing	['demo', 'di', 'tunggangi', 'asing', 'kata', 'prabowo', 'sementara', 'duit', 'negara', 'yg', 'ada', 'di', 'danantara', 'dia', 'kasih', 'asing', 'untuk', 'mengelola', 'mantan', 'koruptor', 'pula', 'ndasmu', 'asing']
simbiosis keknya si kakek mo taruh duit hasil korupnya disini juga lumayan kan dilaundry di danantara	['simbiosis', 'keknya', 'si', 'kakek', 'mo', 'taruh', 'duit', 'hasil', 'korupnya', 'disini', 'juga', 'lumayan', 'kan', 'dilaundry', 'di', 'danantara']

Tabel 5 mengilustrasikan proses tokenisasi, langkah penting dalam pemrosesan bahasa alami yang melibatkan pembagian teks menjadi unit-unit yang lebih kecil dan mudah dikelola yang dikenal sebagai token. Segmentasi ini memungkinkan analisis bahasa yang lebih bernuansa. Misalnya, pada tweet pertama, kalimat "potensi danantara untuk dorong ekonomi nasional sangat besar" diubah menjadi token ['potensi', 'danantara', 'untuk', 'dorong', 'ekonomi', 'nasional', 'sangat', 'besar']. Setiap token mewakili elemen yang bermakna dari teks asli, yang memungkinkan tugas komputasional berikutnya seperti analisis sentimen, pengambilan informasi, dan aplikasi pembelajaran mesin dilakukan dengan lebih efektif.

3.7. Stopword Removal

Stopword Removal merujuk pada kata-kata yang sering muncul dan sering tidak penting dalam teks tertentu yang biasanya dihilangkan selama pemrosesan data. Penghapusan ini berfungsi untuk mengefisienkan data, sehingga meningkatkan efisiensi dan akurasi keseluruhan tugas dan analisis klasifikasi teks [13]. Metodologi penelitian melibatkan pengecualian sistematis kata-kata yang mengandung

kurang dari empat huruf untuk meningkatkan ketepatan analisis.

Tabel 6. Hasil *Stopword Removal*

Twitter	Stopword Removal
['potensi', 'danantara', 'untuk', 'dorong', 'ekonomi', 'nasional', 'sangat', 'besar']	['potensi', 'danantara', 'dorong', 'ekonomi', 'nasional']
['demo', 'di', 'tunggangi', 'asing', 'kata', 'prabowo', 'sementara', 'duit', 'negara', 'yg', 'ada', 'di', 'danantara', 'dia', 'kasih', 'asing', 'untuk', 'mengelola', 'mantan', 'koruptor', 'pula', 'ndasmu', 'asing']	['demo', 'tunggangi', 'asing', 'prabowo', 'duit', 'negara', 'danantara', 'kasih', 'asing', 'mengelola', 'mantan', 'koruptor', 'ndasmu', 'asing']
['simbiosis', 'keknya', 'si', 'kakek', 'mo', 'taruh', 'duit', 'hasil', 'korupnya', 'disini', 'juga', 'lumayan', 'kan', 'dilaundry', 'di', 'danantara']	['simbiosis', 'keknya', 'si', 'kakek', 'mo', 'taruh', 'duit', 'hasil', 'korupnya', 'lumayan', 'dilaundry', 'danantara']

Tabel 6 menyajikan temuan yang diperoleh dari penghapusan kata henti yang umum, termasuk kata hubung dan kata ganti, dengan demikian menyoroti dampak penghapusannya pada keseluruhan analisis data dan proses interpretasi. Misalnya, pada tweet pertama, kata "untuk" dan "sangat" dihapus, sehingga hanya tersisa kata-kata yang lebih relevan seperti "potensi", "danantara", "dorong", dan "ekonomi". Proses yang sama diterapkan pada tweet lainnya, di mana kata-kata seperti "di", "yg", "pula", dan "kan" dihapus, sehingga menghasilkan token yang lebih bersih dan fokus pada kata-kata penting. Stopword removal membantu meningkatkan efisiensi analisis teks dengan menyaring kata-kata yang kurang informatif.

3.8. Stemming

Stemming adalah proses linguistik yang mereduksi kata-kata ke bentuk dasarnya, yang dicontohkan oleh transformasi "bertemu" menjadi "temu." Teknik ini meningkatkan efisiensi pemrosesan teks dengan menyederhanakan variasi kata, sehingga memudahkan pengambilan dan analisis informasi yang lebih baik dalam sentimen analisis [13].

Tabel 7. Hasil *Stemming*

Twitter	Stemming
['potensi', 'danantara', 'dorong', 'ekonomi', 'nasional']	potensi danantara dorong ekonomi nasional
['demo', 'di', 'tunggangi', 'asing', 'kata', 'prabowo', 'sementara', 'duit', 'negara', 'yg', 'ada', 'di', 'danantara', 'dia', 'kasih', 'asing', 'untuk', 'mengelola', 'mantan', 'koruptor', 'pula', 'ndasmu', 'asing']	demo tunggang asing prabowo duit negara danantara kasih asing kelola mantan koruptor ndasmu asing
['simbiosis', 'keknya', 'si', 'kakek', 'mo', 'taruh', 'duit', 'hasil', 'korupnya', 'lumayan', 'dilaundry', 'danantara']	simbiosis kek si kakek mo taruh duit hasil korup lumayan dilaundry danantara

Dari Tabel 7, dapat dilihat hasil dari proses *stemming*, yaitu penyederhanaan kata-kata menjadi bentuk dasar atau akarnya. Misalnya, pada tweet pertama, kata-kata seperti "dorong" dan "nasional" tetap berada dalam bentuk dasarnya, sementara kata "potensi" dan "danantara" tetap seperti aslinya karena sudah dalam bentuk dasar. Pada tweet kedua, kata "mengelola" diubah menjadi "kelola", dan kata "koruptor" disederhanakan menjadi "korup". Proses *stemming* ini bertujuan untuk mengurangi variasi kata yang tidak perlu, sehingga analisis teks dapat lebih fokus pada makna dasar dari kata-kata yang digunakan.

3.9. Future Extraction (Tf-Idf)

TF-IDF, atau Term Frequency-Inverse Document Frequency, adalah metode ekstraksi fitur canggih yang meningkatkan analisis data dengan mengukur frekuensi kata-kata tertentu dalam dokumen tertentu sekaligus menilai kelangkaannya di seluruh kumpulan dokumen yang lebih luas, sehingga memfasilitasi wawasan yang lebih bermakna [5].

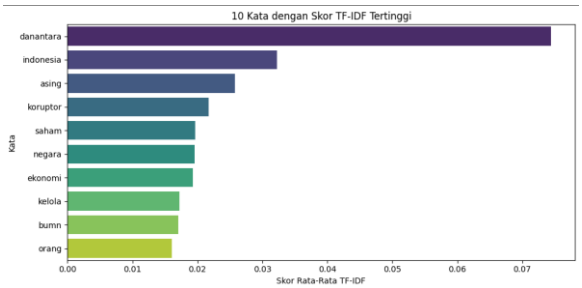
3.10. Support Vector Machine

Support Vector Machine (SVM) adalah teknik klasifikasi canggih yang digunakan dalam pembelajaran mesin yang secara efektif memisahkan data ke dalam kelas-kelas yang berbeda. Hal ini dicapai dengan mengidentifikasi hyperplane optimal, yang diposisikan secara strategis untuk memaksimalkan margin antara titik-titik data terdekat dari setiap kelas, sehingga meningkatkan akurasi dan ketahanan klasifikasi dalam berbagai aplikasi [5].

4. HASIL DAN PEMBAHASAN

4.1. Dataset

Studi ini menggunakan kumpulan data yang dikurasi dengan cermat dan bersumber secara sistematis dari berbagai platform twitter, dengan total 1175 tweet yang dianalisis. Setelah pengumpulan data, dilakukan tahap preprocessing teks pada dataset, yang mencakup pembersihan teks, tokenisasi, dan pelabelan sentimen.



Gambar 1. Frekuensi Kata

Data tersebut mengungkap sepuluh variabel yang paling sering ditemui, menyoroti pola dan tren signifikan dalam kumpulan data yang diamati, antara lain "danantara", "indonesia", "asing", "koruptor", "saham", "negara", "ekonomi", "kelola", "bumi", dan

"orang". Word Cloud berfungsi sebagai alat visualisasi yang efektif, yang memungkinkan representasi sentimen menyeluruh beserta emosi positif dan negatif yang berbeda, sehingga memudahkan pemahaman bernuansa tentang data tekstual dan opini publik.

4.2. Tahap Visualisasi Word Cloud

Pengguna Word Cloud bertujuan untuk merepresentasikan kumpulan data secara visual untuk meningkatkan pemahaman, mengidentifikasi pola, dan memfasilitasi komunikasi data yang efektif. Gambar 2 merupakan hasil visualisasi Word Cloud untuk dataset analisis sentimen "Danantara".

Gambar 2. Word Cloud Sentiment

Dari hasil visualisasi pada Gambar 2, dapat dilihat bahwa kata yang berukuran besar menunjukkan kata-kata yang sering muncul, seperti "danantara", "indonesia", "koruptor", "saham", dan "asing". Istilah-istilah ini mencerminkan topik yang paling sering dibahas oleh pengguna Twitter terkait isu ini, serta memberikan gambaran mengenai fokus sentimen masyarakat terhadap berbagai topik yang berhubungan dengan ekonomi dan politik.

4.3. Tahap Pengujian

Tabel 8. Rancangan Analisis Komputasi

Akurasi	Precision	Recall	F1-Score
79.57%	79.54%	69.62%	71.63%

Penerapan analisis sentimen melalui metode Support Vector Machine (SVM) telah menghasilkan hasil yang patut dicatat, dengan akurasi mencapai 79,57%. Metrik ini menunjukkan bahwa model berhasil mengidentifikasi sentimen yang benar dalam sebagian besar kasus. Precision yang diperoleh sebesar 79,54% menunjukkan bahwa sebagian besar tweet yang diprediksi sebagai sentimen positif atau negatif benar-benar sesuai dengan kategori tersebut. Namun, tingkat Recall kembali sebesar 69,62% menunjukkan tantangan mendasar dalam kemampuan model untuk menangkap semua sentimen positif atau negatif yang relevan yang diungkapkan dalam tweet yang dianalisis. Hal ini menunjukkan bahwa meskipun banyak sentimen diklasifikasikan dengan benar, sejumlah besar diabaikan, yang berpotensi mendistorsi analisis keseluruhan. Lebih jauh, Skor F1 sebesar 71,63% menyoroti keseimbangan antara presisi dan perolehan kembali, yang menekankan perlunya penyempurnaan dalam model untuk meningkatkan kemampuan klasifikasinya yang komprehensif.

5. KESIMPULAN DAN SARAN

Berdasarkan temuan studi ini, jelas bahwa analisis sentimen yang dilakukan pada Program Danantara dalam platform X menggunakan metode Support Vector Machine (SVM) memberikan wawasan yang berharga tentang persepsi publik terhadap program tersebut. Dalam penelitian ini, dilakukan pengumpulan data melalui Twitter yang mencakup berbagai variasi opini masyarakat mengenai Danantara. Setelah melalui tahap preprocessing data, seperti pembersihan, tokenisasi, dan pelabelan sentimen, data dianalisis menggunakan kedua metode algoritma tersebut.

Dari hasil pengujian, Model Support Vector Machine (SVM) menunjukkan kinerja yang baik, mencapai tingkat akurasi 79,57% dalam prediksinya. Metode ini berhasil mengidentifikasi sentimen negatif yang dominan, terkait dengan isu-isu seperti korupsi dan pengelolaan asing yang menjadi perhatian publik. Selain itu, visualisasi data menggunakan Word Cloud juga mengungkapkan kata-kata kunci yang sering muncul, seperti "koruptor", "asing", dan "danantara", yang mengindikasikan fokus utama opini masyarakat terhadap masalah pengelolaan dan investasi negara.

Penelitian ini memberikan kontribusi penting dalam bidang analisis sentimen publik, terutama dalam mengukur persepsi masyarakat terhadap kebijakan publik dan program investasi besar. Temuan ini diharapkan dapat membantu pembuat kebijakan dalam merumuskan strategi komunikasi dan pengelolaan program Danantara yang lebih baik, dengan memperhatikan sentimen dan kekhawatiran masyarakat yang terungkap melalui analisis sentimen berbasis teknologi.

Secara keseluruhan, penelitian ini menegaskan pentingnya penggunaan metode analisis sentimen dalam memahami dinamika opini publik yang dapat mempengaruhi keputusan strategis dan kebijakan nasional. Hasil penelitian ini juga menekankan perlunya pengawasan ketat terhadap pengelolaan dana negara agar terhindar dari potensi penyalahgunaan.

DAFTAR PUSTAKA

[1] Badan Pengelola Investasi Daya Anagata Nusantara, “Daya Anagata Nusantara – Danantara,” 2025-02-24. Accessed: Apr. 12, 2025. [Online]. Available: https://www.danantaraindonesia.com/

[2] A. Indraini, “Apa itu Danantara? Badan Pengelola Investasi yang Baru Diluncurkan Prabowo,” 2025-02-24. Accessed: Apr. 13, 2025. [Online]. Available: https://finance.detik.com/berita-ekonomi-bisnis/d-7792645/apa-itu-danantara-badan-pengelola-investasi-yang-baru-diluncurkan-prabowo

[3] F. Fitriana, E. Utami, and H. Al Fatta, "Analisis Sentimen Opini Terhadap Vaksin Covid - 19 pada Media Sosial Twitter Menggunakan Support Vector Machine dan Naive Bayes," *J. Komtika (Komputasi dan Inform.*, vol. 5, no. 1, pp. 19–25, 2021, doi: 10.31603/komtika.v5i1.5185.

[4] M. Idris, A. Rifai, and K. D. Tania, "Sentiment Analysis of Tokopedia App Reviews using Machine Learning and Word Embeddings," vol. 9, no. 1, pp. 210–219, 2025.

[5] D. Ananda and R. R. Suryono, "Analisis Sentimen Publik Terhadap Pengungsi Rohingya di Indonesia dengan Metode Support Vector Machine dan Naïve Bayes," *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 748, 2024, doi: 10.30865/mib.v8i2.7517.

[6] A. Handayani and I. Zufria, "Analisis Sentimen Terhadap Bakal Capres RI 2024 di Twitter Menggunakan Algoritma SVM," *J. Inf. Syst. Res.*, vol. 5, no. 1, pp. 53–63, 2023, doi: 10.47065/josh.v5i1.4379.

[7] P. Arsi and R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 147, 2021, doi: 10.25126/jtiik.0813944.

[8] A. Hermawan, I. Jowensen, J. Junaedi, and Edy, "Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritma Support Vector Machine," *JST (Jurnal Sains dan Teknol.*, vol. 12, no. 1, pp. 129–137, 2023, doi: 10.23887/jstundiksha.v12i1.52358.

[9] N. M. A. J. Astari, Dewa Gede Hendra Divayana, and Gede Indrawan, "Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier," *J. Sist. dan Inform.*, vol. 15, no. 1, pp. 27–29, 2020, doi: 10.30864/jsi.v15i1.332.

[10] F. A. J. Ayomi and K. E. Dewi, "Analisis Emosi pada Media Sosial Twitter Menggunakan Metode Multinomial Naive Bayes dan Synthetic Minority Oversampling Technique," *Komputa J. Ilm. Komput. dan Inform.*, vol. 12, no. 2, pp. 9–19, 2023, doi: 10.34010/komputa.v12i2.9454.

[11] E. S. Basryah, A. Erfina, and C. Warman, "Analisis Sentimen Aplikasi Dompet Digital Di Era 4 . 0 Pada Masa Pendemi Covid-19 Di Play Store," *SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform.*, pp. 189–196, 2021, [Online]. Available: https://sismatik.nusaputra.ac.id/index.php/sismatik/article/view/28%0Ahttps://sismatik.nusaputra.ac.id/index.php/sismatik/article/download/28/25

[12] N. S. Wardani, A. Prahutama, and P. Kartikasari, "Analisis Sentimen Pemindahan Ibu Kota Negara Dengan Klasifikasi Naïve Bayes Untuk Model Bernoulli Dan Multinomial," *J. Gaussian*, vol. 9, no. 3, pp. 237–246, 2020, doi: 10.14710/j.gauss.v9i3.27963.

[13] M. N. Rahman, *Analisis Performa Penggunaan Stopwords Dan Stemming Dalam Sentimen Analisis Dengan Pendekatan Klasifikasi Naive Bayes*. 2022. [Online]. Available: https://repository.uinjkt.ac.id/dspace/bitstream/123456789/65210/1/MUHAMMAD NURFAJRI RAHMAN-FST.pdf