

PUBLIC SENTIMENT ANALYSIS OF DANANTARA POLICY THROUGH SOCIAL MEDIA X USING SVM AND RANDOM FOREST

ANALISIS SENTIMEN PUBLIK TERHADAP KEBIJAKAN DANANTARA MELALUI MEDIA SOSIAL X DENGAN METODE SVM DAN RANDOM FOREST

Gayus Gregorius Ferdinand Djema¹, Ozzi Suria²

^{1,2}Universitas Mercu Buana Yogyakarta, Sleman, Yogyakarta

Email: anandjema87@gmail.com¹, ozzisuria@mercubuana-yogya.ac.id²

Abstract - This study aims to analyze public sentiment toward the establishment of the Danantara Investment Management Agency (Danantara) through the X social media platform (formerly Twitter) using a machine learning-based text classification approach. While sentiment analysis has been widely applied across various domains, there remains a research gap in examining public responses to new national policies particularly Danantara on platform X. A total of 1,713 tweets were collected using Python-based web scraping via Google Colab during the period from February to June 2025. The research involved data preprocessing, manual sentiment labeling, model training using Support Vector Machine (SVM) and Random Forest algorithms, and performance evaluation using metrics such as accuracy, precision, recall, and F1-score. The classification results show that positive sentiment dominates at 55.6%, while negative sentiment accounts for 44.4%. Random Forest outperformed SVM with an accuracy of 92.36% and an F1-score of 92.19%, compared to SVM's accuracy of 85.45% and F1-score of 87.54%. These findings indicate that Random Forest is more effective in handling short-text public opinion data that is often unstructured. Practically, this study recommends the integration of real-time sentiment monitoring through social media as a strategic tool for policymakers and state-owned enterprises (SOEs) in formulating more responsive and data-driven public policies.

Keywords – Sentiment Analysis, Danantara, Social Media X, SVM, Random Forest.

Abstrak - Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap kebijakan pembentukan Badan Pengelola Investasi Daya Anagata Nusantara (Danantara) melalui media sosial X (sebelumnya Twitter) dengan pendekatan klasifikasi teks berbasis machine learning. Meskipun analisis sentimen telah banyak diterapkan di berbagai bidang, masih terdapat kekosongan penelitian yang secara khusus mengevaluasi respons publik terhadap kebijakan strategis baru seperti Danantara di platform X. Data dikumpulkan menggunakan metode web scraping berbasis Python di Google Colab, dengan total 1.713 tweet yang diperoleh selama periode Februari hingga Juni 2025. Proses penelitian mencakup tahap preprocessing data, pelabelan sentimen, pelatihan model menggunakan algoritma Support Vector Machine (SVM) dan Random Forest, serta evaluasi kinerja model berdasarkan metrik akurasi, precision, recall, dan F1-score. Hasil klasifikasi menunjukkan dominasi sentimen positif sebesar 55,6%, sedangkan sentimen negatif mencapai 44,4%. Random Forest menunjukkan kinerja lebih unggul dengan akurasi 92,36% dan F1-score 92,19%, dibandingkan SVM dengan akurasi 85,45% dan F1-score 87,54%. Temuan ini menunjukkan bahwa Random Forest lebih efektif dalam menangani teks pendek dan opini publik yang tidak terstruktur. Secara praktis, studi ini merekomendasikan agar pembuat kebijakan dan BUMN memanfaatkan pemantauan sentimen berbasis media sosial sebagai alat bantu evaluasi kebijakan publik yang lebih responsif dan berbasis data.

Kata Kunci – Analisis Sentimen, Danantara, Social Media X, SVM, Random Forest.

I. PENDAHULUAN

Media sosial kini berperan penting di era digital. Platform ini memungkinkan publik untuk menyuarakan pendapat, kritik, dan keinginan mereka mengenai berbagai kebijakan pemerintah. Media sosial adalah sarana daring yang memfasilitasi pengguna untuk berinteraksi dan memproduksi konten secara efisien. Keberadaan media sosial telah menjadi sarana bagi masyarakat dalam menyampaikan pandangan serta opini terhadap berbagai isu. Salah satu media sosial yang populer secara global dan sering dimanfaatkan untuk mengekspresikan pendapat publik adalah platform X [1]. Karakteristik platform X yang bersifat terbuka dan cepat dalam penyebaran informasi menjadikannya sumber informasi yang komprehensif untuk mengkaji persepsi dan sikap publik terhadap berbagai isu kebijakan.

Salah satu kebijakan strategis yang mencuat ke ruang publik adalah pembentukan Badan Pengelola Investasi Daya Anagata Nusantara (Danantara). Kebijakan ini secara resmi diluncurkan oleh Presiden Prabowo Subianto pada Februari 2025, telah memunculkan berbagai tanggapan, termasuk kritik dan keraguan dari sejumlah elemen masyarakat. Dari sudut pandang kebijakan, inisiatif ini dilatarbelakangi oleh upaya negara untuk memperkuat kedaulatan ekonomi sebagai langkah strategis menuju percepatan pembangunan ekonomi nasional. Dalam konteks tersebut, pemerintah menekankan pentingnya pengelolaan Badan Usaha Milik Negara (BUMN) yang terencana, akuntabel, transparan, sinergis, dan berkelanjutan. Pendekatan ini bertujuan untuk memastikan partisipasi optimal BUMN dalam mendukung perekonomian nasional yang adaptif terhadap dinamika dan persaingan di tingkat global, baik dalam konteks saat ini maupun di masa yang akan datang [2]. Demi mewujudkan visinya sebagai pengelola investasi unggulan, BPI Danantara menjalankan misi profesional dan berkelanjutan dalam mengelola kekayaan negara, dengan berlandaskan prinsip good governance. Fokusnya mencakup investasi berkualitas di sektor prioritas seperti energi, infrastruktur, dan digitalisasi, serta kemitraan strategis dengan investor asing. Tujuan akhirnya adalah memberikan imbal hasil jangka panjang, mempercepat pembangunan infrastruktur, menciptakan lapangan kerja, dan meningkatkan kesejahteraan masyarakat[3].

Untuk memahami berbagai respon masyarakat terhadap kebijakan Danantara, pendekatan analisis sentimen digunakan untuk mengidentifikasi kecenderungan opini publik dalam bentuk sentimen positif maupun negatif. Metode ini memanfaatkan teknik pembelajaran mesin (machine learning) yang dirancang untuk menganalisis data teks dari media sosial dan mendeteksi pola sentimen yang muncul di dalamnya. Dua algoritma yang sering digunakan dalam tugas klasifikasi teks ini adalah Support Vector Machine (SVM) dan Random Forest, yang dikenal memiliki performa tinggi dalam mengelola data berdimensi besar serta mampu menghasilkan akurasi yang kompetitif dalam konteks analisis sentimen social [4].

Penelitian ini bertujuan untuk mengevaluasi dan membandingkan efektivitas dua algoritma klasifikasi, yakni SVM dan Random Forest, dalam mengidentifikasi serta mengelompokkan sentimen publik terhadap kebijakan Danantara. Evaluasi kinerja dilakukan dengan menggunakan metrik akurasi, precision, recall, dan F1-score sebagai indikator utama. Diharapkan bahwa hasil dari penelitian ini dapat memberikan pemahaman yang lebih mendalam terkait persepsi masyarakat terhadap kebijakan Danantara, sekaligus menjadi referensi yang bernilai bagi para pengambil kebijakan dalam menilai serta mengoptimalkan implementasi program tersebut. Namun, penelitian ini memiliki keterbatasan pada cakupan data yang dikumpulkan hanya dalam rentang waktu tertentu (Februari hingga Juni 2025) serta terbatasnya klasifikasi sentimen yang hanya mencakup dua kategori, yaitu positif dan negatif, tanpa mempertimbangkan kategori netral.

II. SIGNIFIKASI STUDI

A. *Studi Literatur*

Analisis sentimen merupakan bidang ilmu yang berfokus pada pengidentifikasian kecenderungan opini individu atau kelompok, yang umumnya disampaikan melalui komentar atau tulisan. Melalui pendekatan ini, informasi penting mampu diekstraksi dari data tidak terstruktur, terutama opini-opini yang mencerminkan sikap emosional pengguna. Pendapat tersebut kemudian mampu diklasifikasikan menuju kategori sentimen seperti positif atau negatif. Proses pengelompokan ini biasanya dijalankan menggunakan metode klasifikasi teks[5]. Berdasarkan sejumlah studi sebelumnya, algoritma SVM dan Random Forest sama-sama menunjukkan performa klasifikasi yang baik dalam analisis sentimen. Penelitian oleh Syafia et al. mencatat bahwa SVM memiliki akurasi lebih tinggi (96% data latih, 85% data uji) dibandingkan Random Forest (82% dan 80%) dalam klasifikasi komentar YouTube terkait BTS [6]. Studi oleh Septiana dan Alita dalam konteks Pemilu 2024 juga menempatkan SVM di atas Random Forest dengan akurasi 80% dan 78% secara berurutan [7]. Namun, Tangke et al. menunjukkan bahwa Random Forest mencapai akurasi 93%, lebih tinggi dibandingkan SVM (90%) dalam klasifikasi ulasan TikTok berskala besar, menjadikannya pilihan yang lebih stabil terhadap data beragam dan risiko overfitting [8].

Temuan-temuan tersebut menunjukkan bahwa efektivitas SVM dan Random Forest sangat dipengaruhi oleh karakteristik data. Dalam konteks kebijakan publik seperti Danantara yang kompleks dan kaya akan variasi bahasa serta makna tersirat, Random Forest dianggap lebih adaptif karena kemampuannya menangani keragaman struktur bahasa dan mengurangi risiko overfitting. Meski kedua metode ini umum digunakan dalam analisis sentimen, penerapannya untuk mengkaji opini publik terkait kebijakan Danantara di media sosial X masih jarang diteliti. Penelitian ini memberikan manfaat bagi berbagai pemangku kepentingan, khususnya pembuat kebijakan dan pengelola kebijakan Danantara. Kementerian dan lembaga terkait dapat memanfaatkan hasil analisis ini untuk memahami persepsi publik secara real-time dan menyesuaikan strategi komunikasi kebijakan secara lebih efektif. Bagi pihak Danantara, temuan sentimen publik dapat dijadikan indikator awal untuk mengantisipasi potensi resistensi sosial serta meningkatkan transparansi dan keterlibatan publik.

B. *Metodologi Penelitian*

1. Tahapan Penelitian

Tahapan penelitian ini disusun secara bertahap untuk memperoleh hasil analisis sentimen yang akurat terhadap opini masyarakat mengenai kebijakan Danantara melalui platform media sosial X. Secara umum, alur penelitian meliputi proses pengumpulan data, preprocessing data, pelabelan (labeling), pelatihan dan pengujian model klasifikasi, visualisasi data, serta evaluasi kinerja model.

Proses Studi ini diawali dengan melakukan pengumpulan data pada media sosial X menggunakan google colab dan bahasa pemrograman Phyton. Data yang diperoleh akan melalui tahap preprocessing. Selanjutnya data yang telah diproses akan dilabeli dengan dua kategori sentimen yaitu Positif dan Negatif. Data kemudian dibagi menjadi dua bagian yaitu data training dan data testing yang akan diimplementasikan guna menguji performa dari algoritma SVM dan Random Forest. Hasil dari proses klasifikasi kemudian ditampilkan dalam bentuk visualisasi wordcloud yang memiliki peran dalam menunjukkan istilah yang paling dominan atau kerap ditemukan dalam dataset. Sebagai tahap akhir, kinerja metode klasifikasi yang digunakan dianalisis dengan menerapkan metrik evaluasi seperti akurasi dan F1-score, yang bertujuan untuk mengukur tingkat keberhasilan model dalam mengklasifikasikan data secara tepat[9].

2. Pengumpulan Data

Data yang digunakan pada penelitian ini dikumpulkan dari media sosial X dengan menggunakan kata kunci “Danantara” dan “Daya Anagata Nusantara” melalui scraping di Google Collab dengan menggunakan bahasa pemrograman Phyton [10]. Dari proses tersebut berhasil diperoleh sebanyak 1.713 tweet pengguna yang kemudian dijadikan sebagai dataset utama dalam proses analisis sentimen terhadap persepsi masyarakat atas terbentuknya Danantara.

3. Preprocessing Data

Pada tahap ini dataset akan diolah dengan berbagai tahapan guna memastikan kualitas dri dataset yang ada, agar tidak terjadinya gangguan pada saat melakukan proses traning data. Preprocessing data dilakukan dengan mengikuti tahapan sebagai berikut :

- a. Melakukan pemeriksaan pada dataset apakah terdapat missing value dan data data yang duplikat, jika terdapat data yang abnormal seperti itu maka dapat dilakukan tindakan yang sesuai, seperti menghapus beberapa hal yang tidak diperlukan seperti tanda baca, emoticon, URL, dan sebagiannya[11].
- b. *Case Folding* merupakan tahap normalisasi teks yang bertujuan untuk mengonversi seluruh huruf kapital menjadi huruf kecil, sehingga dapat menyamakan representasi kata yang sebenarnya identik namun ditulis dengan format huruf yang berbeda. Proses ini penting untuk memastikan konsistensi dalam analisis data teks[12].
- c. Normalisasi merupakan tahap dalam preprocessing teks yang bertujuan untuk menyelaraskan bentuk kata dengan standar bahasa Yang tepat, dengan cara mengganti kata tidak baku, singkatan, atau istilah informal menjadi bentuk baku. Proses ini penting untuk memastikan bahwa sistem pemrosesan bahasa alami dapat mengenali dan memperlakukan kata-kata yang memiliki kesamaan secara konsisten, yang pada akhirnya memperbaiki tingkat ketepatan dalam klasifikasi sentimen[13].
- d. Tokenisasi adalah komponen dari proses pengolahan teks yang berfungsi untuk memisahkan suatu kalimat atau teks yang dibagi ke dalam segmen teks yang disebut token, yaitu unit terkecil seperti kata atau frasa yang akan dianalisis lebih lanjut[14].
- e. Filtering merupakan tahap penyederhanaan data teks dengan cara menghapus istilah umum yang mempunyai makna lemah atau tidak signifikan, seperti kata hubung atau kata sambung. Kata-kata tersebut tidak memberikan kontribusi berarti terhadap proses analisis, sehingga penghapusannya bertujuan untuk meningkatkan efisiensi dan fokus model dalam analisis sentimen[15][22].
- f. Dalam tahap preprocessing data, stemming digunakan untuk menstandarkan kata dengan cara menghilangkan imbuhan, sehingga setiap kata direpresentasikan dalam bentuk dasarnya untuk meningkatkan efektivitas analisis[16].

4. Labeling

Labeling adalah proses pengelompokan data teks menjadi dua kategori klasifikasi, yakni sentimen positif dan negatif, berdasarkan nilai sentimen yang diperoleh. Penilaian ini mengacu pada kamus sentimen yang telah terstandarisasi, di mana kata-kata dalam teks dievaluasi dan diberi skor sesuai dengan kecenderungan emosionalnya, mengacu pada referensi yang telah banyak digunakan dalam penelitian sebelumnya [17].

5. Klasifikasi Metode

SVM dan Random Forest merupakan dua algoritma yang secara luas diterapkan dalam analisis text mining, terutama untuk keperluan klasifikasi teks karena kemampuannya menangani data berukuran besar dan kompleks. SVM mengoptimalkan pemisahan antar kelas melalui hyperplane dan fungsi kernel untuk menangani data non-linier. Di sisi lain, Random Forest menggabungkan

banyak decision tree melalui sistem voting untuk hasil yang lebih stabil dan tahan terhadap overfitting. Keduanya banyak diterapkan dalam analisis berbasis teks seperti sentimen, topik, dan opini. [18].

6. Visualisasi Wordcloud

Word Cloud adalah salah satu bentuk visualisasi data teks yang menyajikan kata-kata dari sebuah dokumen ukuran setiap kata pada visualisasi menunjukkan intensitas kemunculannya dalam dokumen, berdasarkan distribusi frekuensi. Kata yang memiliki frekuensi tinggi akan ditampilkan dengan ukuran yang lebih besar, sedangkan kata-kata yang jarang muncul akan ditampilkan dengan ukuran yang lebih kecil. Pendekatan visual ini memberikan cara yang intuitif dan cepat untuk mengamati kata-kata yang paling menonjol dalam kumpulan teks, sehingga memudahkan pemahaman terhadap konten dominan dalam dokumen tersebut[19][23].

7. Evaluasi Metode

Data yang telah dibersihkan dikelompokkan menjadi dua sentimen dan dibagi menjadi data latih dan uji. Model SVM dan Random Forest diuji pada data uji, lalu dievaluasi berdasarkan confusion matrix serta metrik kinerja seperti akurasi, presisi, dan recall untuk mengetahui efektivitas klasifikasi sentiment [24].

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pengumpulan data dilakukan dengan cara web scraping, yaitu metode yang diterapkan untuk mengekstraksi informasi dari halaman-halaman situs web, termasuk media sosial X. Dalam Studi ini, data tweet terhadap kebijakan Danantara pada sosial media X diperoleh melalui proses web scraping yang diimplementasikan menggunakan google collab.



Gambar 1. Pengambilan Data

Gambar 1 mengilustrasikan urutan proses pengambilan data yang diterapkan dalam penelitian ini. Tahapan ini dimulai dari identifikasi istilah penting yang berkaitan langsung dengan topik kajian, seperti “Danantara” dan istilah terkait lainnya. Selanjutnya, proses pengumpulan data dilakukan dengan menjalankan skrip Python di Colab untuk melakukan web scraping terhadap media sosial X. Hasil dari proses ini berupa Kumpulan tweet yang disimpan ke dalam file dengan format CSV.

B. Preprocessing Data

Preprocessing memegang peranan penting dalam meningkatkan kualitas data sebelum diterapkan pada model klasifikasi. Dalam penelitian ini, tahapan preprocessing yang mencakup cleaning, case folding, normalisasi, tokenisasi, penghapusan stopwords, dan stemming berkontribusi signifikan terhadap akurasi model dalam menganalisis sentimen. Proses ini tidak hanya merapikan dan menstandarkan struktur teks, tetapi juga membantu mengurangi gangguan dari elemen-elemen tidak relevan seperti URL, simbol, maupun emotikon. Dengan demikian, preprocessing tidak dapat dipandang sebagai tahapan pendahuluan semata, melainkan sebagai komponen esensial yang memengaruhi keberhasilan klasifikasi, khususnya dalam konteks analisis teks pendek dan tidak terstruktur seperti opini masyarakat di media sosial. Hasil preprocessing ditampilkan pada table 1.

Tabel 1. Preprocessing

Tahapan Preprocessing	Hasil
Content	Apakah ini akan menjadi ladang KORUPSI lagi ? Semoga tidakmenambah jumlah Koruptor Di Indonesia. 👉 https://t.co/ywIOxQDalv
Cleaning & Case Folding	apakah ini akan menjadi ladang korupsi lagi semoga tidak menambah jumlah koruptor di Indonesia
Normalisasi	apakah ini akan jadi ladang korupsi lagi semoga tidak tambah jumlah koruptor di Indonesia
Tokenisasi	['apakah', 'ini', 'akan', 'jadi', 'ladang', 'korupsi', 'lagi', 'semoga', 'tidak', 'tambah', 'jumlah', 'koruptor', 'di', 'indonesia']
Filtering	['jadi', 'ladang', 'korupsi', 'lagi', 'tambah', 'jumlah', 'koruptor', 'indonesia']
Stemming	['jadi', 'ladang', 'korupsi', 'lagi', 'tambah', 'jumlah', 'koruptor', 'indonesia']

C. Labeling

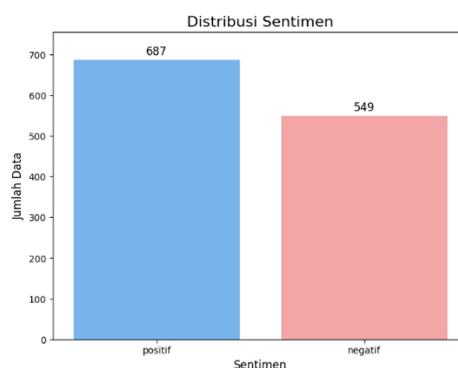
Setelah tahap preprocessing selesai dilakukan, proses dilanjutkan dengan pemberian label pada data. Data akan dilabeli dengan sentimen positif dan negatif. Tahap labeling ini dilakukan dengan melabeli data atau tweet yang sebelumnya telah di cleaning. Hasil dari labelling ditampilkan pada tabel 2.

Tabel 2. Labeling

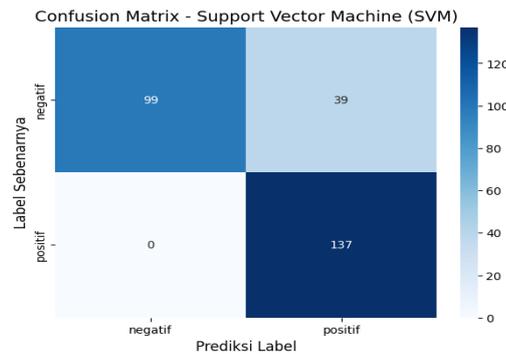
Tweet	Label
presiden prabowo subianto badan kelola investasi daya anagata nusantara investasi proyekproyek strategis sektor energi ta	Positif
perintah tugas badan kelola investasi daya anagata nusantara garap proyek hilir proyekproyek risiko mati investor ekonomis	Negatif
badan kelola investasi daya anagata nusantara indonesia badan kelola investasi strategis konsolidasi optimal investasi perintah dukung tumbuh ekonomi nasional	Positif

D. Klasifikasi Metode

Setelah preprocessing, sebanyak 1.236 tweet dibagi ke dalam dua kelas sentimen: positif dan negatif. Model yang dihasilkan kemudian disimpan sebagai bagian dari proses pembentukan model klasifikasi yang akan digunakan pada tahap evaluasi selanjutnya. Hasil dari implementasi menunjukkan bahwa sentimen positif lebih dominan dengan persentase sebesar 55.6% yang terdiri dari 687 tweet. Sementara itu sentimen negatif tercatat sebesar 44.4 % yang terdiri dari 549 tweet. Temuan ini mengindikasikan bahwa sebagian besar masyarakat memberikan reaksi yang positif terhadap kebijakan yang akan dijalankan. Gambar 3 menunjukkan hasil dari implementasi.

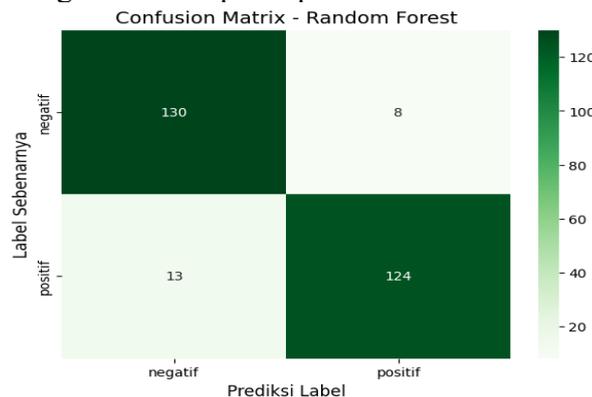


Gambar 3. Diagram Hasil Analisis Sentimen



Gambar 4. Confusion Matrix SVM

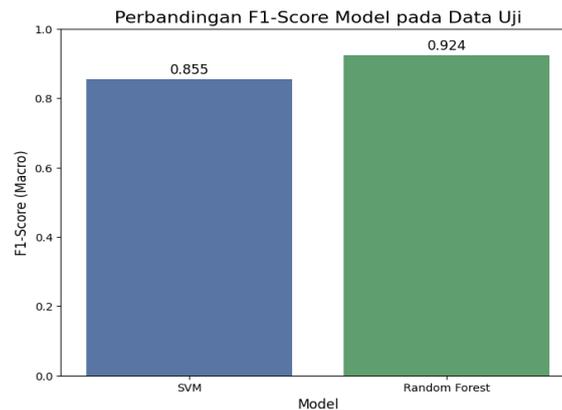
Gambar 4 merupakan gambar confusion matrix yang memperlihatkan hasil klasifikasi sentimen menggunakan algoritma SVM terhadap data opini masyarakat terkait kebijakan Danantara. Berdasarkan hasil tersebut, sebanyak 99 data berlabel negatif berhasil diprediksi dengan benar (true negative), sementara 39 data negatif diklasifikasikan keliru sebagai sentimen positif (false positive). Di sisi lain, seluruh data berlabel positif berhasil diprediksi secara tepat oleh model, yakni sebanyak 137 data (true positive), dan tidak ditemukan kesalahan klasifikasi pada data positif (false negative = 0). Dengan demikian, model mencapai tingkat akurasi sebesar 85,45%, dengan nilai precision untuk sentimen positif mencapai 77,84% dan nilai recall sebesar 100%. Recall 100% berarti model SVM tidak melewatkan satu pun data positif. Semua data yang seharusnya diklasifikasikan sebagai positif berhasil terklasifikasi dengan benar. Namun, bukan berarti model ini sempurna, karena terdapat False Positive sebanyak 39, yang berdampak pada turunnya precision. Sementara itu, nilai F1-score yang dihasilkan berada pada angka 87,54%. Hasil ini sejalan dengan penelitian sebelumnya yang dilakukan oleh Syafia et al. [6] serta Septiana dan Alita [7], yang menunjukkan bahwa algoritma SVM unggul dalam mengenali sentimen positif, namun cenderung kurang presisi dalam mengklasifikasikan sentimen negative. Oleh karena itu, meskipun efektif dalam mendeteksi sentimen positif, model ini masih perlu dioptimalkan untuk meningkatkan ketepatan pada klasifikasi sentimen negatif.



Gambar 5. Confusion Matrix Random Forest

Gambar 5 merupakan gambar confusion matrix yang menggambarkan hasil klasifikasi model Random Forest terhadap data sentimen masyarakat mengenai kebijakan Danantara. Berdasarkan hasil evaluasi, model berhasil mengklasifikasikan 130 data negatif secara benar (true negative) dan 124 data positif secara akurat (true positive). Sementara itu, terdapat 8 data negatif yang salah diklasifikasikan sebagai positif (false positive) dan 13 data positif yang keliru diklasifikasikan sebagai negatif (false negative). Dari distribusi tersebut, diperoleh tingkat akurasi sebesar 92,36%, dengan nilai presisi dan recall masing-masing mencapai 93,93% dan 90,51%. Nilai F1-score yang dihasilkan sebesar 92,19% menunjukkan keseimbangan yang baik antara presisi dan recall. Hasil ini sejalan dengan studi yang dilakukan oleh Tangke et al. [8], yang mengungkapkan bahwa algoritma Random Forest menunjukkan performa lebih baik dibandingkan SVM, dengan tingkat

akurasi yang mencapai 93%. Perbandingan dari kinerja antar model SVM dan Random Forest dapat dilihat dari gambar 6 dibawah ini.



Gambar 6. Perbandingan Model

IV. KESIMPULAN

Hasil penelitian menunjukkan bahwa algoritma Random Forest memberikan kinerja yang lebih optimal dibandingkan SVM dalam mengklasifikasikan sentimen masyarakat terhadap kebijakan Danantara di media sosial X. Dari total 1.236 tweet yang dianalisis, sentimen positif mencakup 55,6%, sementara sentimen negatif sebesar 44,4%. Random Forest berhasil mencapai akurasi 92,36% dengan F1-score 92,19%, lebih tinggi dibandingkan SVM yang mencatat akurasi 85,45% dan F1-score 87,54%. Hal ini mencerminkan keunggulan Random Forest dalam menghasilkan klasifikasi yang lebih stabil dan seimbang. Proporsi sentimen positif yang lebih tinggi menunjukkan adanya dukungan dari publik terhadap kebijakan tersebut, namun kehadiran sentimen negatif yang cukup signifikan menandakan perlunya perhatian terhadap aspek transparansi dan komunikasi publik. Penelitian ini memiliki keterbatasan dalam hal cakupan data yang terbatas pada periode Februari hingga Juni 2025, serta hanya menggunakan dua kategori sentimen tanpa mempertimbangkan kategori netral. Untuk pengembangan ke depan, disarankan penggunaan pendekatan multi-label, teknik representasi teks lanjutan seperti word embedding, serta perluasan data ke platform lain agar hasil analisis lebih komprehensif. Secara aplikatif, hasil penelitian ini dapat dimanfaatkan oleh instansi pemerintah dan BUMN untuk mengembangkan sistem pemantauan opini publik secara real-time yang terintegrasi dengan media sosial, guna mendeteksi persepsi masyarakat terhadap kebijakan secara cepat dan akurat. Sistem ini dapat digunakan sebagai alat deteksi dini terhadap potensi resistensi atau krisis kepercayaan publik.

REFERENSI

- [1] F. F. Rachman and S. Pramana, "Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter," *Heal. Inf. Manag. J.*, vol. 8, no. 2, pp. 100–109, 2020, [Online]. Available: <https://inohim.esaunggul.ac.id/index.php/INO/article/view/223/175>
- [2] Vina Hardyana Infantri and Retno Meilani, "Analisis Yuridis Pembentukan Badan Pengelola Investasi Daya Anagata Nusantara (BPI Danantara)," *J. Ris. Rumpun Ilmu Sos. Polit. dan Hum.*, vol. 4, no. 1, pp. 378–391, May 2025, doi: 10.55606/jurrish.v4i1.5132.
- [3] O. Defilania and W. Silalahi, "Badan Pengelola Investasi Daya Anagata Nusantara : Peluang Dan Tantangan Dalam Reformasi Ekonomi Indonesia Daya Anagata Nusantara Investment Management Agency : Opportunities and Challenges in Indonesian Economic Reform," no. April, pp. 7125–7134, 2025.
- [4] K. Purifyregalia, K. Umam, N. Cahyo, H. Wibowo, and M. R. Handayani, "Detecting Fake Reviews in E-Commerce : A Case Study on Shopee Using Support Vector Machine and Random Forest," vol.

- 9, no. 3, 2025.
- [5] B. Ramadhani and R. R. Suryono, “Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse,” *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 714, 2024, doi: 10.30865/mib.v8i2.7458.
- [6] A. N. Syafia, M. F. Hidayattullah, and W. Suteddy, “Studi Komparasi Algoritma SVM Dan Random Forest Pada Analisis Sentimen Komentar Youtube BTS,” *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, pp. 207–212, 2023, doi: 10.30591/jpit.v8i3.5064.
- [7] I. Septiana and D. Alita, “Perbandingan Random Forest dan SVM dalam Analisis Sentimen Quick Count Pemilu 2024,” vol. 9, no. 3, pp. 224–233, 2024, doi: 10.30591/jpit.v9i3.6640.
- [8] R. Tangke *et al.*, “Analisis Sentimen Aplikasi Tiktok Menggunakan Algoritma Support Vector Machine (Svm) Dan Random Forest,” vol. XIII, no. 2, pp. 53–62, 2024.
- [9] D. A. Fitri, “Komparasi algoritma random forest classifier dan support vector machine untuk sentimen masyarakat terhadap pinjaman online di media sosial,” vol. 9, no. 4, pp. 2018–2029, 2024.
- [10] L. Rofiqi and M. Akbar, “Analisis Sentimen Terkait RUU Perampasan Aset dengan Support Vector Machine,” *JEKIN - J. Tek. Inform.*, vol. 4, no. 3, pp. 529–538, 2024, doi: 10.58794/jekin.v4i3.824.
- [11] E. M. Pusung and I. N. Dewi, “Optimasi RoBERTa dengan Hyperparameter Tuning untuk Deteksi Emosi berbasis Teks,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 10, no. 3, pp. 240–248, 2025.
- [12] M. Joefitra Zaqy, L. Marlina, and R. F. Wijaya, “Analysis of Indonesian Netizen Sentiment on Platform X Regarding the Arrival of Refugees in Indonesia Using the Multinomial Naive Bayes Method,” *Sinkron*, vol. 8, no. 3, pp. 1945–1952, 2024, doi: 10.33395/sinkron.v8i3.13940.
- [13] A. Agung, A. Daniswara, I. Kadek, and D. Nuryana, “Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru,” *J. Informatics Comput. Sci.*, vol. 05, pp. 97–100, 2023.
- [14] D. S. Nurrochmah, N. Rahaningsih, R. D. Dana, and C. L. Rohmat, “Penerapan Algoritma Naive Bayes dalam Analisis Sentimen Ulasan Aplikasi KitaLulus di Google Play Store,” *J. Inform. Terpadu*, vol. 11, no. 1, pp. 1–11, 2025, doi: 10.54914/jit.v11i1.1544.
- [15] A. Santosa, I. Purnamasari, and Mayasari Rini, “Pengaruh Stopword Removal dan Stemming Terhadap Performa Klasifikasi Teks Komentar Kebijakan New Normal Menggunakan Algoritma LSTM,” *J. Sains Komput. Inform.*, vol. 6, pp. 81–93, 2022.
- [16] K. Lukman and S. Novianto, “Komparasi Algoritma Naïve Bayes dan SVM untuk Identifikasi Cyberbullying Selebriti di Media Sosial Twitter,” pp. 970–981, 2025, doi: 10.33364/algoritma/v.22-1.2196.
- [17] D. Ayu, P. Wulandari, G. W. Wardhana, I. D. Gede, and P. Wirata, “Sentiment Analysis of the Free Lunch Program by Prabowo Gibran Using the Naive Bayes Classifier,” vol. 4, no. 3, pp. 173–181, 2025.
- [18] T. Antaka, J. Laga, and N. Widjiyati, “JTIM : Jurnal Teknologi Informasi dan Multimedia Comparison of Social Media Between Tiktok and Instagram to Detect Negative Content Using Natural Language Processing Method,” vol. 7, no. 3, pp. 433–439, 2025.
- [19] M. Galih Pradana, “Penggunaan Fitur Wordcloud Dan Document Term Matrix Dalam Text Mining,” *J. Ilm. Informatika*, vol. 08, no. 01, pp. 38–43, 2020.
- [20] M. Fathoni, “Word Cloud Analysis for The Urgency of The Role of Islamic Schools Headmaster in Merdeka Curriculum Implementation Analisis Word Cloud terhadap Urgensi Peran Kepala Madrasah dalam Implementasi Kurikulum Merdeka,” *J. 12 Waiheru*, vol. 10, no. 1, 2024.
- [21] N. Hendrastuty, A. Rahman Isnain, and A. Yanti Rahmadhani, “Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine,” *J. Inform. J. Pengemb. IT*, vol. 6, no. 3, pp. 150–155, 2021.
- [22] S. Setyabudi and E. Aryanny, “Sentiment Analysis of Lazada Marketplace User Ratings with Naïve Bayes and Support Vector Machine Methods,” *INOVTEK Polbeng - Seri Inform.*, vol. 10, no. 1, pp. 422–433, 2025, doi: 10.35314/sww8cg21.
- [23] M. Iqrom, M. Afdal, Rice Novita, Medyantiwi Rahmawita, and Tengku Khairil Ahsyar, “Sentiment Analysis of Gojek, Grab, Maxim Applications Using Support Vector Machine Algorithm,” *INOVTEK Polbeng - Seri Inform.*, vol. 10, no. 1, pp. 237–248, 2025, doi: 10.35314/52fycr56.
- [24] R. Randy Suryono, “Sentiment Analysis of the Influence of the Korean Wave in Indonesia Using the Naive Bayes Method and Support Vector Machine Analisis Sentimen Pengaruh Korean Wave Di Indonesia Dengan Metode Naive Bayes Dan Support Vector Machine,” *J. Inovtek Polbeng - Seri Inform.*, vol. 10, no. 1, pp. 308–319, 2025.