Jurnal TRANSFORMATIKA
Vol.24, No.1, Juli 2026, pp. 104 - 121
P-ISSN: 1693-3656, E-ISSN: 2460-6731
https://journals.usm.ac.id/index.php/transformatika/

Analisis Sentimen Publik Pada Media Sosial X Terhadap Pembentukan Danantara Menggunakan Metode *Support Vector Machine*

Aqilah Dz. Tulimau[1], Emerensye. S. Y. Pandie[2], Clarissa Elfira Amos Pah[3*]
[1]Program Studi Ilmu Komputer, Fakultas Sains dan Teknik, Universitas Nusa Cendana
Jalan Adi Sucipto Penfui, (0380) 881580, e-mail: tulimauaqilah@gmail.com
[2]Program Studi Ilmu Komputer, Fakultas Sains dan Teknik, Universitas Nusa Cendana
Jalan Adi Sucipto Penfui, (0380) 881580, e-mail: emerensyepandie@staf.undana.ac.id
[3*]Program Studi Ilmu Komputer, Fakultas Sains dan Teknik, Universitas Nusa Cendana
Jalan Adi Sucipto Penfui, (0380) 881580, e-mail: clarissaelfira@staf.undana.ac.id

ARTICLE INFO

History of the article:
Received 22 April 2026
Received in revised form 23 Juli 2026
Accepted 23 Juli 2026
Available online 12 Agustus 2026

Keywords:

Analisis Sentimen; Danantara; TF-IDF; *Lexicon InSet*; *Support Vector Machine*

*** Correspondence:**

Telepon :
+62 81339232250

E-mail:
clarissaelfira@staf.undana.ac.id

ABSTRACT

Pembentukan Danantara di Indonesia menimbulkan berbagai respons masyarakat, sehingga diperlukan analisis sentimen untuk memahami persepsi publik terhadap kebijakan tersebut. Penelitian ini bertujuan untuk menganalisis sentimen publik terkait pembentukan Danantara pada platform X ke dalam kelas sentimen positif, netral, dan negatif. Kebaruan penelitian ini terletak pada perbandingan performa empat *kernel Support Vector Machine*, yaitu Linear, *Polynomial*, *Radial Basis Function*, dan *Sigmoid*, serta evaluasi sentimen sebelum dan setelah peresmian Danantara. Data diproses melalui tahapan *text preprocessing*, pelabelan sentimen berbasis *InSet Lexicon*, pembobotan TF-IDF, serta klasifikasi menggunakan metode SVM. Pengujian dilakukan menggunakan pembagian data 70:30, 80:20, dan 90:10. Hasil penelitian menunjukkan bahwa rasio 90:10 memberikan performa terbaik, dengan *Polynomial* memperoleh akurasi tertinggi sebesar 91,66%. Selain itu, sentimen publik didominasi oleh kelas negatif sebesar 65,9% dan 60,0% pada periode sebelum dan setelah peresmian Danantara. Hasil ini menunjukkan bahwa SVM terbukti efektif untuk klasifikasi sentimen dan dapat digunakan sebagai evaluasi kebijakan serta strategi komunikasi publik.

PENDAHULUAN

Investasi merupakan pilar penting dalam mendorong pertumbuhan ekonomi suatu negara, sehingga pemerintah Indonesia melakukan suatu langkah strategis yaitu membentuk badan pengelola investasi yang berfungsi untuk mengoptimalkan dan mengelola aset negara. Pada tanggal 24 Februari 2025, Presiden Republik Indonesia, Prabowo Subianto, secara resmi meluncurkan Badan Pengelola Investasi (BPI) Daya Anagata Nusantara (Danantara) [1]. Pembentukan Danantara didasarkan pada rancangan UU perubahan ketiga atas UU Nomor19 Tahun 2003 tentang BUMN dan Peraturan Pemerintah No 10 Tahun 2025, yang mengatur tentang organisasi dan tata kelola Danantara[2].

Pembentukan BPI Danantara menimbulkan beragam reaksi dari masyarakat. Sebagian pihak menolak karena aset negara seolah dipertaruhkan dalam investasi, apabila terjadi kegagalan dapat menyebabkan bencana ekonomi dan bertambahnya utang negara. Sebaliknya, sebagian pihak mendukung karena berdampak positif bagi masyarakat yakni dapat membuka lapangan pekerjaan baru jika Danantara dapat dikelola dengan baik [3]. Berbagai reaksi dari masyarakat tersebut banyak dituangkan melalui berbagai platform media sosial, salah satunya adalah platform X. Berdasarkan survei dari *We Are Social* yang dilakukan pada bulan Januari tahun 2025, menunjukkan bahwa pengguna X di Indonesia cukup tinggi, di mana sekitar 57,5% dari jumlah populasi atau sekitar 105– 110 juta pengguna pada masyarakat usia 16 hingga 64 tahun yang tercatat sebagai pengguna aktif platform X [4]. Tingginya jumlah pengguna tersebut, menunjukkan bahwa X menjadi salah satu media sosial dengan tingkat partisipasi publik yang tinggi sebagai sarana penyampaian opini. Selain itu, X memiliki keunggulan dalam menampilkan komentar pengguna secara *real-time* serta dilengkapi dengan fitur *hashtag* sehingga mempermudah pengelompokan topik yang berkaitan dengan isu tertentu [5]. Salah satu isu yang banyak dibahas pada platform X adalah pembentukan Danantara, hal ini ditunjukkan dengan tingginya aktivitas pengguna pada penggunaan *hastag* Danantara yang mencapai hampir 90 ribu opini. Dengan demikian, analisis sentimen diperlukan untuk mengidentifikasi dan mengelompokkan kecenderungan opini masyarakat secara sistematis, sehingga dapat diketahui dominasi sentimen yang muncul, baik positif, negatif, maupun netral terhadap pembentukan Danantara. Dalam melakukan analisis sentimen, terdapat berbagai metode yang dapat digunakan, salah satunya adalah Metode *Support Vector Machine* (SVM).

Metode SVM telah terbukti efektif dalam melakukan klasifikasi sentimen teks karena kemampuannya menangani data berdimensi tinggi dengan menemukan *hyperplane* yang optimal [6]. Meskipun *hyperplane* digunakan untuk memisahkan dua kelas, metode SVM tetap dapat digunakan untuk proses klasifikasi *multiclass* (positif, negatif, dan netral) pada klasifikasi sentimen publik melalui pendekatan One-vs-All [7]. One-vs-All bekerja dengan membagi permasalahan *multiclass* menjadi sejumlah permasalahan biner, yaitu sebanyak jumlah kelas yang ada, kemudian setiap kelas akan dibandingkan dengan semua kelas lainnya [8]. Dengan kemampuan yang dimiliki, beberapa penelitian terdahulu telah menerapkan SVM untuk analisis sentimen terhadap berbagai konteks isu publik, seperti pelayanan publik, evaluasi kebijakan pemerintah, ulasan aplikasi digital, hingga opini terkait produk dan seni. Penelitian oleh [9] menghasilkan akurasi 87% pada penerapan *kernel* linear dan *sigmoid* SVM untuk klasifikasi sentimen vaksin Covid-19, penelitian [10] untuk opini terhadap program Kartu Prakerja mencapai akurasi sebesar 98,67% pada *kernel* linear, analisis sentimen pemilihan presiden 2024 oleh [11] menghasilkan rata-rata akurasi sebesar 95,18% menggunakan keempat *kernel* (linear 95,45%, *sigmoid* 95,30%, *polynomial* 95,05%, dan *Radial Basis Function* (RBF) 94,92%), [12] melakukan analisis sentimen mengenai Permendikbud dan kekerasan seksual di kampus yang mendapatkan akurasi sebesar 69,15%, dan analisis sentimen opini masyarakat terhadap film horor Indonesia oleh [13] dengan akurasi sebesar 82,51%.

Penelitian terkait sentimen masyarakat terhadap Danantara juga telah dilakukan sepanjang tahun 2025 hingga 2026 yang ditunjukkan pada tabel 1.

Tabel 1. Gap penelitian terdahulu dan kontribusi penelitian

Penelitian	**Metode, *Kernel*, Akurasi, Pelabelan, Interval Waktu Data Sentimen**	***Gap* Penelitian Terdahulu**	**Kontribusi Penelitian Ini (Menjawab *Gap* Penelitian Terdahulu)**
[14]	- Metode Klasifikasi: SVM - *Kernel*: RBF - Akurasi 68,7% - Metode Pelabelan: *Lexicon* - Periode Data: Februari – Maret 2025	Hanya terbatas pada satu jenis *kernel* (RBF) dan tidak ada perbandingan sentimen sebelum dan setelah penetapan Danantara. Data yang digunakan hanya setelah pembentukan Danantara.	Membandingkan 4 *kernel* SVM dan melakukan analisis sentimen sebelum vs sesudah peresmian. Performa akurasi yang lebih baik.
[15]	- Metode Klasifikasi:- - Metode Pelabelan: *InSet Lexicon* - Akurasi: Tidak ada akurasi - Periode Data: 23 Februari – 7 Maret 2025	Tidak menggunakan algoritma klasifikasi untuk generalisasi data (terbatas pada pelabelan) dan tidak ada perbandingan sentimen sebelum dan setelah penetapan Danantara. Data yang digunakan hanya setelah pembentukan Danantara. Selain itu, tidak ada metode evaluasi yang dapat membuktikan performa dari pelabelan.	Menggabungkan *Lexicon* dengan SVM serta evaluasi pergeseran opini. Dengan menggunakan metode klasifikasi seperti SVM, pelabelan *lexicon* dapat dievaluasi keakuratannya.
[16]	- Metode Klasifikasi: SVM - *Kernel*: Tidak disebutkan - Akurasi: Tidak ada akurasi: 79,57% - Metode Pelabelan: Otomatis (*Deberta Python*) - Periode Data: Januari – April 2025	Tidak membandingkan performa berbagai jenis *kernel* dan tidak mengevaluasi fase kebijakan secara spesifik.	Melakukan komparasi performa 4 *kernel* dan analisis distribusi dua periode waktu berbeda, serta menghasilkan performa model yang lebih akurat.
[17]	- Metode Klasifikasi: *Hybrid* (IndoRoBERTa + SVM) - *Kernel*: RBF - Akurasi: 98%	Tidak menyebutkan secara spesifik rentang waktu pengumpulan data sehingga tidak dapat menentukan apakah data yang diambil merupakan fase sebelum atau sesudah penetapan Danatara.	Menggunakan pelabelan *InSet Lexicon* yang divalidasi serta optimasi 4 *kernel* SVM. Analisis yang lebih mendalam dengan pembedaan

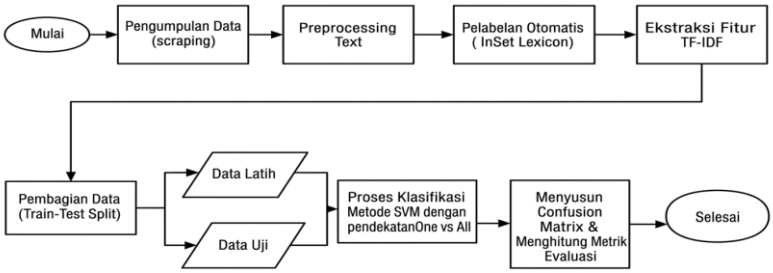
Penelitian	Metode, *Kernel*, Akurasi, Pelabelan, Interval Waktu Data Sentimen	*Gap* Penelitian Terdahulu	Kontribusi Penelitian Ini (Menjawab *Gap* Penelitian Terdahulu)
	- Metode Pelabelan: Pseudo-labeling (IndoRoBERTa) - Periode Data: -		fase sebelum dan setelah penetapan Danantara.
[18]	- Metode Klasifikasi: *Naïve Bayes* dan KNN - Akurasi: 68,89% - Metode Pelabelan: Manual - Periode Data: Februari – April 2025	Menghasilkan akurasi yang cukup rendah dibandingkan dengan penggunaan SVM untuk kasus yang sama. Selain itu, pelabelan dengan metode manual yang tidak dapat dipastikan validitasnya.	Menggunakan SVM dengan perbandingan *kernel* untuk mencari model klasifikasi paling optimal.
[19]	- Metode Klasifikasi: SVM - *Kernel*: RBF - Akurasi: 87% - Metode Pelabelan: Otomatis (*Library Lexicon*) - Periode Data: 24 Februari – 4 April 2025	Evaluasi hanya menggunakan satu *kernel* (RBF) dan hanya menggunakan data setelah penetapan Danantara, sehingga tidak dapat membandingkan sentimen sebelum dan setelah penetapan Danantara.	Implementasi 4 jenis *kernel* dan menganalisis pergerakan opini sebelum dan sesudah kebijakan. Penelitian ini juga mendapatkan akurasi lebih tinggi.
[20]	- Metode Klasifikasi: *Naive Bayes* - Akurasi: 79,9% - Metode Pelabelan: Manual dengan *matching keyword* - Periode Data: 1 Januari 2024 – 12 Desember 2025	Memiliki akurasi yang lebih rendah dari rata-rata penelitian menggunakan SVM dan tidak membedakan fase waktu kebijakan secara mendalam.	Melakukan eksperimen perbandingan *kernel* SVM dan mengukur pergeseran sentimen. Penelitian ini juga mendapatkan akurasi lebih tinggi.
[21]	- Metode Klasifikasi: IndoBERT - Akurasi: 97,71% - Metode Pelabelan: Manual - Periode Data: 1 Januari – 22 April 2025	Menawarkan kebaruan dengan menggunakan metode *deep learning* (LLM), namun pelabelan masih manual, sehingga tidak dapat dipastikan validitasnya. Tidak ada perbandingan sentimen sebelum dan setelah penetapan Danantara	Menawarkan model SVM yang lebih efisien secara komputasi tetapi tetap menghasilkan akurasi yang tinggi.
[22]	- Metode Klasifikasi: *Hybrid* (K-*Means* + SVM	Pelabelan otomatis K-*Means* berisiko bias tanpa validasi linguistik dan periode data	Menerapkan pelabelan *InSet Lexicon* dan melakukan analisis

Penelitian	Metode, *Kernel*, Akurasi, Pelabelan, Interval Waktu Data Sentimen	*Gap* Penelitian Terdahulu	Kontribusi Penelitian Ini (Menjawab *Gap* Penelitian Terdahulu)
	- *Kernel*: *Linear* - Akurasi: 98% - Pelabelan: - Periode Data: Tidak menyebutkan interval waktu yang spesifik	tidak disebutkan secara spesifik.	komparatif temporal secara mendalam.

Berdasarkan tinjauan penelitian terdahulu, dapat disimpulkan bahwa penelitian ini secara efektif mengisi celah dari peneliti sebelumnya. Sebagian besar penelitian terdahulu cenderung hanya menggunakan satu jenis *kernel* SVM (terutama RBF atau Linear), atau hanya fokus pada algoritma tertentu tanpa mempertimbangkan pergeseran sentimen. Penelitian ini menjawab kekurangan tersebut dengan memberikan perbandingan performa empat *kernel* SVM (Linear, *Polynomial*, RBF, dan *Sigmoid*) secara mendalam, di mana ditemukan bahwa *kernel Polynomial* menghasilkan akurasi tertinggi sebesar 91,66%. Selain itu, penelitian ini memberikan kontribusi dengan melakukan analisis komparatif (sebelum dan sesudah peresmian) pada isu Danantara, yang memungkinkan identifikasi pergeseran sentimen publik akibat peristiwa kebijakan secara spesifik. Dengan demikian, penelitian ini tidak hanya memberikan kontribusi teknis pada optimasi model, tetapi juga kontribusi praktis dalam evaluasi kebijakan pemerintah.

METODE PENELITIAN

Alur penelitian ditunjukkan oleh Gambar 1, yang diawali dengan proses pengumpulan data ulasan dengan cara *scrapping*. Data yang diperoleh kemudian diproses melalui tahap *text preprocessing*, pelabelan menggunakan *lexicon inset*, ekstraksi fitur TF-IDF, pembagian data dengan menggunakan 3 skenario, selanjutnya dilakukan klasifikasi dengan metode SVM, dan diakhiri dengan tahap pengujian kinerja model.



Gambar 1. Alur penelitian

1. Pengumpulan data

Data dikumpulkan menggunakan teknik *scrapping* dari *platform* X dengan kata kunci terkait "Danantara" dengan rentang waktu 1 Desember 2024 hingga 31 Oktober 2025. Dari hasil *scrapping* diperoleh sebanyak 88.917 ulasan. Data tersebut kemudian digunakan dalam proses klasifikasi yang dikelompokkan ke dalam tiga kelas sentimen yaitu positif, negatif, dan netral.

2. *Text Processing*

Pada tahap *Text preprocessing*, data ulasan dibersihkan dan ditransformasikan agar informasi yang terkandung didalamnya dapat diproses secara efektif pada tahap selanjutnya. Proses

dilakukan melalui beberapa tahapan yaitu *cleaning*, *case folding*, normalisasi, *tokenizing*, *stemming*, *convert negation*, *stopword removal* [23][24].

a. *Cleaning* yaitu proses membersihkan data teks dari karakter-karakter yang tidak relevan (seperti *emoji* dan simbol) yang dapat mempengaruhi kinerja model dalam proses klasifikasi.
b. *Case Folding* adalah proses menyeragamkan seluruh huruf dalam teks menjadi *lowercase*.
c. *Tokenizing* adalah proses memecahkan teks menjadi unit-unit kata yang disebut token.
d. Normalisasi adalah proses mengubah semua kata menjadi bentuk baku.
e. *Stemming* adalah proses mengubah kata ke bentuk dasar.
f. *Convert Negation* adalah proses untuk menangani kata-kata negasi seperti "belum" dan sejenisnya, agar makna kalimat dapat dipahami degan tepat oleh sistem.
g. *Stopword Removal* adalah proses penghapusan kata umum yang tidak mengandung makna, seperti kata "dari", "yang", "ke", dan sejenisnya.

3. Pelabelan dengan *Inset-Lexicon*

InSet (Indonesian Sentiment) Lexicon adalah pendekatan dalam analisis sentimen berbasis kamus kata untuk menentukan sentimen sebuah teks. Kamus kata tersebut berisi daftar kata dengan label polaritas tertentu serta dilengkapi bobot atau skor dari setiap kata [25]. *InSet lexicon* berfungsi memberikan label otomatis pada data latih sebelum masuk ke dalam tahap pelatihan model [26]. Dalam proses pelabelan, setiap teks diberikan skor polaritas berdasarkan jumlah kata yang terdapat dalam kamu *lexicon*. Berdasarkan nilai skor tersebut, data diklasifikasikan ke dalam tiga kelas sentimen yaitu positif, negatif, dan netral. Rumus *InSet Lexicon* ditunjukkan pada persamaan 1.

$$polarity_{score} = \sum_{i=1}^{n} score\ (i) \quad (1)$$

Di mana $Polarity_{Score}$ adalah total nilai polaritas dari suatu kalimat (*post*/komentar pada sosial media), n adalah jumlah kata yang memiliki nilai polaritas, dan $score(i)$ adalah nilai Bobot nilai sentimen kata ke-i yang terdaftar dalam kamus *InSet* (rentang -5 hingga +5)[27].

Penggunaan *InSet Lexicon* dipertimbangkan dalam penelitian ini karena merupakan metode pelabelan yang cukup unggul dalam mengklasifikasikan sentimen dan kerap diintegrasikan dengan SVM. Hal ini dapat terlihat dari penelitian terdahulu yang telah melakukan perbandingan antara *InSet Lexicon* dengan metode pelabelan lainnya. Penelitian [28] melakukan perbandingan antara metode *InSet Lexicon* dengan *Ratting-Based labeling* yang mendapati akurasi *InSet Lexicon* lebih tinggi sebesar 89.7%. Selanjutnya kedua penelitian [29] dan [30] yang membandingkan pelabelan menggunakan *Valence Aware Dictionary and sEntiment Reasoner* (VADER) *Lexicon* dan *InSet Lexicon* menghasilkan akurasi yang lebih tinggi pada *InSet Lexicon*. Selain keunggulan performa akurasi, *InSet Lexicon* banyak diterapkan dalam analisis sentimen Indonesia karena dikembangkan khusus dari karakteristik media sosial karena kamus ini dibangun dengan mengumpulkan kata-kata langsung dari data *mikroblog* (X) berbahasa Indonesia[15]. Kamus *InSet* disusun secara manual pada penelitian [27] dan telah dievaluasi performanya secara langsung dengan membandingkan beberapa kamus *lexicon*. *InSet* mencapai akurasi tertinggi sebesar 65,78%, mengungguli berbagai leksikon populer yang diterjemahkan ke bahasa Indonesia seperti *SentiWordNet*, Liu *Lexicon*, dan AFINN, serta melampaui performa Vania *Lexicon* (salah satu *Lexicon* Indonesia lainnya) [27]. Kamus ini terdiri dari 3.609 kata positif dan 6.609 kata negatif dengan rentang skor polaritas antara -5 hingga +5. Keterlibatan manusia dalam pembobotan kamus *InSet* memastikan bahwa makna kata yang masuk ke dalam kamus benar-benar merepresentasikan emosi atau sentimen yang sesuai dengan konteks budaya dan linguistik Indonesia.

4. Ektraksi fitur dengan TF-IDF

Ekstraksi fitur dilakukan dengan menggunakan metode *Frequency–Inverse Document Frequency* untuk merepresentasikan teks dalam bentuk vektor numerik sehingga dapat diolah oleh algoritma pembelajaran mesin [31]. Tahapan dalam TF-IDF meliputi perhitungan TF (*Term Frequency*), IDF (*Inverse Document Frequency*), bobot TF-IDF, dan Normalisasi TF-IDF, untuk menghasilkan representasi fitur yang konsisten[8][17]. Persamaan 2 digunakan untuk menghitung TF yang menunjukkan tingkat kemunculan suatu kata dalam dokumen.

$$TF_{t,d} = \frac{f_{t,d}}{n_d} \tag{2}$$

Di mana $TF_{t,d}$ adalah *Term frequency* dari *term* t pada dokumen d, kemudian $f_{t,d}$ adalah jumlah kemunculan kata t pada dokumen d, dan n_d adalah Jumlah kata dalam dokumen d [8]. Persamaan 3 digunakan untuk menghitung IDF yang menunjukkan seberapa penting suatu *term* muncul dalam semua dokumen.

$$IDF = \log \frac{N}{df_t} \tag{3}$$

Di mana $IDFt$ adalah *Invers documen frequency* untuk *term* t, kemudian N adalah total seluruh dokumen, dan dft adalah jumlah dokumen yang mengandung *term* t [8]. Persamaan 4 digunakan untuk menghitung nilai bobot TF-IDF.

$$W_{t,d} = TF_{t,d} \times IDF_t \tag{4}$$

Di mana $W_{t,d}$ adalah bobot TF-IDF pada *term* t dalam dokumen d. Persamaan 5 digunakan untuk menghitung normalisasi TF-IDF, di mana normalisasi TF- IDF adalah proses menyesuaikan nilai vektor TF-IDF agar memiliki skala yang konsisten [8].

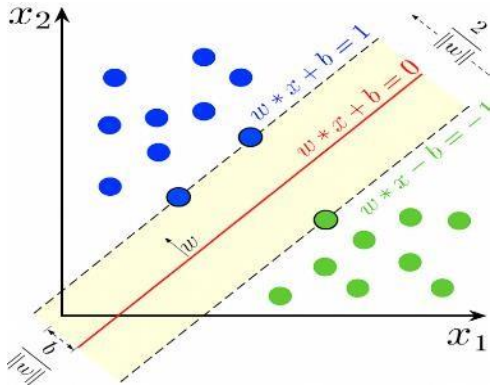
$$W_{(t,d)_{norm}} = \frac{W_{t,d}}{\sqrt{\sum W_{t,d}^2}} \tag{5}$$

5. Pembagian data

Data yang telah diproses melalui tahapan *text preprocessing*, pelabelan *lexicon-inset*, dan pembobotan TF-IDF, selanjutnya akan dibagi menjadi data latih dan data uji menggunakan tiga skenario pembagian, yaitu 70% data latih dan 30% data uji, 80% data latih dan 20% data uji, kemudian 90% data latih dan 10% data uji. Pembagian ini dilakukan untuk mengevaluasi pengaruh proporsi data latih terhadap kinerja model klasifikasi.

6. Klasifikasi dengan Support Vector Machine

Proses klasifikasi dilakukan menggunakan metode *Support Vector Machine* (SVM) dengan pendekatan *One-vs-All* untuk menangani klasifikasi *multi-class*. SVM adalah merupakan salah satu algoritma dalam *supervised learning* yang digunakan untuk menganalisis data, mengenali pola, serta melakukan tugas klasifikasi maupun regresi. Dalam proses klasifikasi, SVM bekerja dengan menentukan *hyperplane* atau batas pemisah (*decision boundary*) yang optimal untuk memisahkan satu kelas dengan kelas lainnya. *Hyperplane* terbaik diperoleh dengan memaksimalkan margin, yaitu jarak antara *hyperplane* dengan titik data terdekat (*support vector*) pada masing-masing kelas[32]. Secara umum, komponen utama dalam SVM meliputi *hyperplane* optimal, kelas positif, kelas negatif, serta margin [6]. Gambar 2 menunjukkan ilustrasi dari metode SVM [33].



Gambar 2. Ilustrasi *support vector machine*

SVM menyediakan fungsi *kernel* yang dapat digunakan dalam proses klasifikasi data, dapat berupa pola *linear* maupun non-*linear*. Jenis-jenis *kernel* SVM adalah sebagai berikut [34][35]:

a. *Linear*
Kernel linear digunakan untuk memetakan data secara langsung melalui operasi *dot product*, sebagaimana ditunjukkan pada persamaan 6.

$$K(xp, xq) = (xp \ . \ xq) \tag{6}$$

b. *Polynomial*
Kernel polynomial digunakan untuk memetakan data ke ruang fitur berdimensi tinggi dengan mempertimbangkan derajat tertentu, sebagaimana ditunjukkan pada persamaan 7.

$$K(xp, xq) = (xp.xq + c)^d \tag{7}$$

c. RBF
RBF adalah *kernel* yang digunakan dalam menangani data *non-linear* berbasis jarak antar vektor fitur, sebagaimana ditunjukkan pada persamaan 8.

$$K(xp, xq) = \exp(-\gamma||xp - xq\ ||^2) \tag{8}$$

d. *Sigmoid*
Kernel sigmoid menggunakan parameter gamma dan konstanta untuk menghasilkan transformasi non-*linear*, sebagaimana yang ditunjukkan pasa persamaan 9.

$$K(xp, xq) = tanh(\alpha\, xp \ . \ xq + c) \tag{9}$$

Di mana $K(xp, xq)$ adalah fungsi *kernel*, c adalah konstanta fungsi *kernel*, d adalah derajat *polynomial*, γ adalah parameter *kurva gaussian*, *tanh* adalah fungsi *hiperbolik tangensial*, dan α adalah konstanta penskalaan. Adapun parameter SVM yang digunakan dalam penelitian ini adalah sebagai berikut:

a. Parameter c = 1.0 untuk mengaktifkan *inhomogeneus* pada kernel *polynomial* sehingga model tetap mempertimbangkan fitur berderajat lebih rendah (fitur *linear* asli), bukan hanya berderajat d=2. Pada sebagian literatur nilai 1 langsung digunakan dalam perumusan seperti pada [36][37]. Sementara nilai 1.0 pada c *kernel sigmoid* berfungsi menjadi bias untuk menentukan *threshold* suatu data akan diidentifikasi sebagai sentimen positif atau negatif oleh fungsi *tanh*.
b. *Kernel polynomial* menggunakan *degree* (d) = 2 (kuadratik) seperti pada [38] yang juga menggunakan *degree*=2 dan menghasilkan akurasi terbaik sebesar 99,4% pada kasus analisis sentimen.

c. *Kernel* RBF dan *Sigmoid* menggunakan parameter gamma = *scale* yang secara dinamis menyesuaikan karakteristik data dengan lebih fleksibel terhadap skala data. *Scale* akan memperkecil nilai gamma secara otomatis agar model tidak terlalu sensitif terhadap perbedaan angka tersebut dan mencegah *overfitting*. *Scale* mempertahankan nilai yang mempertimbangkan jumlah fitur dan varians data [39].

Proses klasifikasi dengan SVM menggunakan pendekatan *One-vs-All* karena menangani klasifikasi *multiclass* yaitu kelas sentimen positif, negatif, dan netral. Cara penerapan *One-vs-All* untuk menentukan prediksi data uji adalah dengan membuat 3 buah model biner yaitu model biner 1 (Positif vs Bukan Positif), model biner 2 (Negatif vs Bukan Negatif), model biner 3 (Netral vs Bukan Netral). Selanjutnya, untuk menentukan kelas prediksi akhir dari masing-masing data uji, dilakukan dengan melakukan pemilihan nilai maksimum dari setiap model biner. Dalam penelitian ini, proses klasifikasi dengan SVM akan menggunakan 4 jenis *kernel*, yaitu *kernel linear*, *polynomial*, RBF, dan *Sigmoid*. Penggunaan keempat *kernel* tersebut bertujuan untuk mengetahui *kernel* dengan performa terbaik. Tahapan dalam melakukan klasifikasi dengan metode SVM adalah sebagai berikut[37][40]:

a. Melakukan perhitungan fungsi *kernel* dengan menggunakan persamaan 6 sampai 9.
b. Menghitung matriks *hessian* dengan menggunakan persamaan 10.

$$H_{ij} = y_i y_j (K(x_p, x_q) + \lambda^2) \quad untuk\ i, j = 1, \dots, n \tag{10}$$

Di mana H_{ij} adalah elemen matriks *hessian* pada baris-i dan kolom-j dan y_i, yj adalah label kelas data ke-i dan ke-j.

c. Mencari nilai *support vector* dengan metode *sequential learning*.
   1) Menghitung *error rate* dengan menggunakan persamaan 11.

$$E_p = \alpha \sum_{j=1}^{n} H_{pj} \tag{11}$$

Di mana E_p adalah nilai *error* dan n adalah jumlah data latih.

   2) Mengitung nilai $\delta\alpha i$ menggunakan persamaan 12.

$$\delta\alpha_p = \min\{\max[\, \gamma\, (1 - E_p), -\alpha]\, , C - \alpha\} \tag{12}$$

Di mana $\delta\alpha_p$ adalah nilai *delta alpha* untuk dokumen ke-p.

   3) Menghitung nilai α_i dengan menggunakan persamaan 13.

$$\alpha_p{}^{baru} = \alpha^{sebelumnya} + \delta\alpha_p \tag{13}$$

Di mana $\alpha_p{}^{baru}$ adalah Nilai *alpha* yang diperbarui.

d. Menghitung nilai bias menggunakan persamaan 14.

$$b = -\frac{1}{2}[w^+ + w^-] \tag{14}$$

Di mana w^+ adalah bobot kelas positif dan w^- adalah bobot kelas negatif.

e. Melakukan tahap pengujian yaitu setiap data uji (testing) akan dibandingkan (dihitung

DOI : https://doi.org/10.26623/transformatika.v24i1.14327

kernel-nya) dengan semua data latih, di mana perhitungannya menggunakan persamaan 15.

$$K(x_{test}, x_{train}) = (x_{test} \cdot x_{train}) \quad (15)$$

f. Menerapkan fungsi klasifikasi untuk mengklasifikasikan data baru berdasarkan model SVM yang telah dilatih, sebagaimana ditunjukkan pada persamaan 16.

$$f(x_k) = sign(\sum_{i=1}^{n} \alpha^{baru}\, y\, K(x_{test}, x_{train}) + b) \quad (16)$$

Di mana $f(x_k)$ adalah fungsi keputusan pada data uji ke-k.

7. Pemetaan *Confusion Matrix*

Confusion matrix adalah matriks yang berisi jumlah prediksi benar dan salah dari model klasifikasi, kemudian menjadi dasar perhitungan berbagai metrik evaluasi seperti akurasi, presisi, *recall*, dan F1-*score* [38]. Struktur umum *confusion matrix* untuk tiga kelas ditujukkan pada Tabel 2.

Tabel 2. Struktur umum *confusion matrix*

	Prediksi Positif	**Prediksi Netral**	**Prediksi Negatif**
Aktual Positif	*True* Positif	*False* Netral (*Positive*)	*False* Negatif (Positif)
Aktual Netral	*False* Positif (Netral)	*True* Netral	*False* Negatif (Netral)
Aktual Negatif	*False* Positif (Negatif)	*False* Netral (Negatif)	*True* Negatif

Berdasarkan *confusion mat*rix pada Tabel 2, dapat dijadikan dasar perhitungan nilai akurasi, presisi, *recall*, dan f1-*score*. Definisi dan rumus perhitungan masing-masing metrik evaluasi disajikan sebagai berikut[38]:

a. Akurasi

Akurasi merupakan metrik yang digunakan untuk mengukur kinerja klasifikasi berdasarkan tingkat keakuratan dari hasil prediksi, yang dilihat dari rasio prediksi benar (positif, netral, dan negatif). Persamaan 17 digunakan untuk menghitung nilai akurasi.

$$Akurasi = \frac{True\ Positif + True\ Netral + True\ Negatif}{Total\ Data} \quad (17)$$

b. Presisi

Presisi yaitu jumlah data prediksi positif yang benar terhadap semua data yang dilabeli positif oleh model. Perhitungan nilai presisi dilakukan menggunakan persamaan 18, 19, dan 20.

$$Presisi_{positif} = \frac{True\ Positif}{True\ Positif + False\ Positif\ (Netral) + False\ Positif(Negatif)} \quad (18)$$

$$Presisi_{netral} = \frac{True\ Netral}{True\ Netral + False\ Netral\ (Positif) + False\ \ Netral(Negatif)} \quad (19)$$

$$Presisi_{negatif} = \frac{True\ Negatif}{True\ Negatif + False\ Negatif(Positif) + False\ Negatif(Netral)} \quad (20)$$

c. Recall

Recall merupakan ukuran yang menunjukkan proporsi data positif yang berhasil diprediksi dengan benar oleh model dari semua data aktual positif. Perhitungan nilai *recall* dilakukan menggunakan persamaan 21, 22, dan 23.

$$Recall_{positif} = \frac{True\ Positif}{True\ Positif + False\ Netral\ (Positif) + False\ Negatif (Positif)} \quad (21)$$

$$Recall_{netral} = \frac{True\ Netral}{True\ Netral + False\ Positif\ (Netral) + False\ Negatif\ (Netral)} \quad (22)$$

$$Recall_{negatif} = \frac{True\ Negatif}{True\ Negatif + False\ Positif\ (Negatif) + False\ Netral\ (Negatif)} \quad (23)$$

d. F1-*Score*

F1-*Score* merupakan nilai rata-rata untuk mengukur performa model dengan mempertimbangkan keseimbangan danatara *presisi* dan *recall* yang dihitung menggunakan persamaan 24, 25, dan 26.

$$F_1 - Score_{positif} = 2\left(\frac{Presisi_{positif} \times Recall_{positif}}{Presisi_{positif} + Recall_{positif}}\right) \quad (24)$$

$$F_1 - Score_{netral} = 2\left(\frac{Presisi_{netral} \times Recall_{netral}}{Presisi_{netral} + Recall_{netral}}\right) \quad (25)$$

$$F_1 - Score_{negatif} = 2\left(\frac{Presisi_{negatif} \times Recall_{negatif}}{Presisi_{negatif} + Recall_{negatif}}\right) \quad (26)$$

HASIL

Pada Data penelitian diperoleh dari *platform* X dengan menerapkan teknik *scraping* yang dimulai dari tanggal 1 Desember 2024 hingga 31 Oktober 2025, dengan jumlah sebanyak 88.917 data ulasan yang berkaitan dengan pembentukan Danantara. Setelah melalui *tahapan text preprocessing*, jumlah data mengalami pengurangan menjadi 74.568, akibat penghapusan data duplikat. Setelah semua teks dibersihkan, langkah berikutnya yaitu melakukan pelabelan dengan *inset lexicon*, setiap teks ulasan diklasifikasikan ke dalam 3 kategori yaitu positif, negatif, dan netral. Gambar 3 menunjukkan beberapa data teks yang telah melalui semua tahap *text preprocessing* dan telah diberi label sentimen.

stopword_removal_tokens	inset_score	skor_per_kata	label_sentimen
[logika, sungsang, danantara, hasil, uang, sum...	-3	logika [0] sungsang [0] danantara [0] hasil [3...	negatif
[danantara, kontan, reaksi, negatif]	-2	danantara [0] kontan [3] reaksi [-5] negatif [0]	negatif
[bukti, nyata, badan usaha milik negara, rugi,...	-14	bukti [-3] nyata [-2] badan usaha milik negara...	negatif
[logika, sungsang, bpi, danantara, hasil, uang...	3	logika [0] sungsang [0] bpi [0] danantara [0] ...	positif
[gila, tidak puas, gaji, segitu, bayar, negara...	-15	gila [-5] tidak puas [-3] gaji [0] segitu [-3]...	negatif

Gambar 3. Hasil *text preprocessing* dan pelabelan sentimen

Tahap selanjutnya yaitu melakukan proses ekstraksi fitur TF-IDF, yang menghasilkan sebanyak 35.562 fitur kata unik yang digunakan sebagai representasi data dalam proses klasifikasi. Kemudian, melakukan pembagian data latih dan data uji dengan menggunakan 3 skenario pembagian data yaitu 70:30, 80:20, dan 90:10. Proses klasifikasi menggunakan metode SVM dengan pendekatan *One-vs-All* untuk menangani tiga kelas sentimen.

Tahap terakhir adalah melakukan perhitungan metrik evaluasi yaitu akurasi, presisi, *recall*, dan f1-*score*, Tabel 3 menunjukkan hasil evaluasi kinerja dari 4 jenis *kernel* SVM pada 3 skenario pembagian data.

Tabel 3. Hasil evaluasi kinerja model

Jenis *Kernel*	Pembagian Data	Kelas	Presisi	*Recall*	F1-*Score*	Akurasi
Linear	70:30	Positif	0.90	0.88	0.89	90.26
		Netral	0.82	0.64	0.72	
		Negatif	0.92	0.97	0.94	

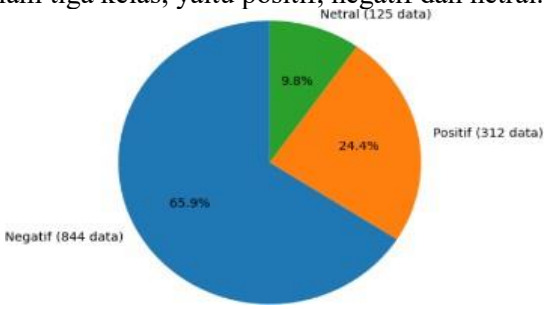
DOI : https://doi.org/10.26623/transformatika.v24i1.14327

Jenis *Kernel*	Pembagian Data	Kelas	Presisi	*Recall*	F1-*Score*	Akurasi
	80:20	Positif	0.90	0.89	0.89	90.25%
		Netral	0.81	0.63	0.71	
		Negatif	0.92	0.96	0.94	
	90:10	Positif	0.91	0.90	0.90	**90.79%**
		Netral	0.82	0.65	0.72	
		Negatif	0.92	0.97	0.94	
Polynomial	70:30	Positif	0.91	0.90	0.91	91.39%
		Netral	0.80	0.69	0.74	
		Negatif	0.93	0.96	0.95	
	80:20	Positif	0.91	0.90	0.91	91.39%
		Netral	0.79	0.68	0.73	
		Negatif	0.93	0.96	0.95	
	90:10	Positif	0.92	0.91	0.91	**91.66%**
		Netral	0.80	0.70	0.75	
		Negatif	0.93	0.96	0.95	
RBF	70:30	Positif				
		Netral	0.90	0.46	0.61	
		Negatif	0.86	0.98	0.92	
	80:20	Positif	0.91	0.84	0.87	87.99%
		Netral	0.88	0.47	0.62	
		Negatif	0.87	0.98	0.92	
	90:10	Positif	0.91	0.85	0.88	**88.68%**
		Netral	0.88	0.51	0.65	
		Negatif	0.88	0.98	0.93	
Sigmoid	70:30	Positif	0.82	0.72	0.77	80.60%
		Netral	0.62	0.30	0.40	
		Negatif	0.82	0.95	0.88	
	80:20	Positif	0.82	0.73	0.78	81.31%
		Netral	0.65	0.37	0.47	
		Negatif	0.83	0.94	0.88	
	90:10	Positif	0.83	0.74	0.79	**82.11%**
		Netral	0.65	0.41	0.50	
		Negatif	0.84	0.94	0.88	

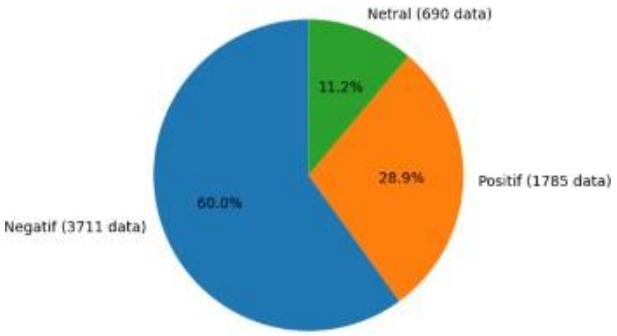
Berdasarkan Tabel 3, nilai akurasi tertinggi diperoleh pada *kernel polynomial* dengan pembagian data 90:10 yaitu 91,66%. Sementara itu, *kernel linear* dengan pembagian data 90:10 yaitu sebesar 90,79%, *kernel* RBF memiliki akurasi tertinggi yaitu 88,68% pada pembagian data 90:10, *kernel sigmoid* juga menghasilkan akurasi tertinggi pada pembagian data 90:10 yaitu sebesar 82.11. selanjutnya, ditinjau berdasarkan metrik evaluasi per kelas sentimen, nilai presisi, *recall*, dan f1-score pada kelas negatif cenderung lebih tinggi dibandingkan kelas positif dan netral pada sebagian besar *kernel* dan skenario pembagian data. Pada *kernel polynomial* dengan pembagian data 90:10, kelas negatif menghasilkan nilai presisi 0,93, *recall* 0,96, dan f1-*score* 0,95. Hasil yang konsisten terdapat pada kernel *linear* pembagian data 70:30 yaitu memiliki nilai presisi sebesar 0,92, *recall* 0,97, dan f1-*score* 0,94. Selain itu, kelas positif juga memiliki nilai presisi, *recall*, dan f1-*score* yang relatif tinggi pada beberapa skenario. Misalnya pada *kernel polynomial* pembagian data 80:20, kelas positif menghasilkan nilai presisi sebesar 0.91, *recall* 0.90, dan f1-*score* 0.91. Sebaliknya, kelas netral memiliki nilai presisi, *recall*, dan f1-*score* yang lebih rendah dibandingkan

kelas lainnya. Pada *kernel polynomial* pembagian data 90:10, kelas netral hanya memperoleh presisi sebesar 0.80, *recall* sebesar 0.91, dan f1-*score* sebesar 0.75. Penurunan lebih signifikan terlihat pada *kernel* RBF pembagian data 70:30, di mana nilai *recall* kelas netral hanya mencapai 0.46 dan *kernel* RBF pembagian data 80:20 menghasilkan *recall* sebesar 0.46, serta pada *kernel sigmoid* pembagian data 70:30 menghasilkan *recall* terendah sebesar 0.30.

Berikutnya adalah analisis distribusi sentimen yang disajikan untuk memberikan gambaran dan mengamati pola kecenderungan opini publik pada dua periode waktu yaitu sebelum peresmian dan setelah peresmian BPI Danantara. Analisis ini dilakukan secara deskriptif berdasarkan hasil klasifikasi sentimen ke dalam tiga kelas, yaitu positif, negatif dan netral.



Gambar 4. Hasil analisis sentimen sebelum peresmian Danantara



Gambar 5. Hasil analisis sentimen setelah peresmian Danantara

Berdasarkan Gambar 4 dan Gambar 5, terlihat sentimen negatif mendominasi pada periode sebelum dan setelah peresmian. Pada periode sebelum peresmian, kelas negatif menghasilkan data sebanyak 844 (65.9%), kelas positif sebanyak 312 data (24,4%), dan kelas netral sebanyak 125 data (9.8%). Pada periode setelah peresmian, kelas negatif menghasilkan data sebanyak 3711 (60.0%), kelas positif sebanyak 1785 data (28,9%), dan kelas netral sebanyak 690 data (11.12%).

PEMBAHASAN

Hasil penelitian menunjukkan bahwa metode SVM mampu melakukan klasifikasi sentimen dengan performa yang baik pada seluruh skenario pembagian data. Berdasarkan hasil pengujian, *kernel polynomial* menghasilkan akurasi tertinggi sebesar 91,66% pada pembagian data 90:10. Hasil tersebut menunjukkan bahwa kernel *polynomial* mampu memberikan performa yang lebih baik dibandingkan *kernel linear*, RBF, dan *sigmoid*. Hal ini dikarenakan *kernel polynomial* mampu memetakkan data ke ruang fitur berdimensi tinggi sehingga pola non-*linear* pada teks dapat dipisahkan dengan lebih optimal. Kemampuan tersebut menyebabkan *kernel polynomial* lebih efektif dalam mengenal pola sentimen dibandingkan *kernel* lainnya. Pada isu kebijakan ekonomi makro seperti Danantara, opini publik cenderung menggunakan kombinasi kata yang kompleks

(misalnya, perpaduan istilah teknis investasi dengan istilah tidak formal bernada kritik). *Kernel Polynomial* mampu menangkap interaksi antar kata tersebut dengan lebih optimal dibandingkan *kernel Linear* yang hanya mencari batas keputusan garis lurus, atau *kernel* RBF dan *Sigmoid* yang dalam beberapa kasus teks media sosial yang padat justru cenderung mengalami *underfitting* atau terlalu sensitif terhadap parameter skala. Hasil penelitian ini sejalan dengan penelitian yang dilakukan oleh [41], yang membandingkan *kernel linear* dan non-*linear* untuk performa klasifikasi sentimen terhadap e-*commerce* di Indonesia pada komentar (*post*) Twitter (X). Penelitian tersebut juga mengangkat permasalahan yang sama dengan penelitian ini, yakni penggunaan bahasa tidak formal pada sosial media membutuhkan metode klasifikasi yang dapat beradaptasi pada dimensi yang lebih kompleks. Hasilnya yang menunjukkan bahwa *kernel polynomial* dari kelompok *kernel* non-*linear* menghasilkan performa klasifikasi yang tinggi secara konsisten dibandingkan dengan RBF dan *Sigmoid* dengan akurasi rata-rata lebih dari 90% dan akurasi tertinggi mencapai 93%. Kesamaan hasil tersebut memperkuat bahwa *kernel polynomial* memiliki kemampuan yang baik dalam menangani klasifikasi data teks, khususnya data sentimen media sosial yang memiliki dimensi fitur tinggi.

Berdasarkan hasil evaluasi per kelas sentimen, kelas negatif memperoleh nilai presisi, *recall*, dan f1-*score* yang relatif tinggi dibandingkan kelas positif dan netral pada sebagian besar *kernel* dan skenario pembagian data. Tingginya performa kelas negatif menunjukkan bahwa model mampu mengenali pola kata yang berkaitan dengan sentimen negatif secara konsisten. Kondisi ini disebabkan oleh karakteristik teks sentimen negatif yang cenderung memiliki kata-kata dengan polaritas yang lebih kuat dan spesifik dibandingkan kelas lainnya. Selain itu, jumlah data kelas negatif yang lebih banyak dibandingkan kelas lainnya juga memungkinkan model mempelajari pola kelas negatif dengan lebih optimal. Sementara itu, kelas netral menunjukkan nilai *recall* dan f1-*score* yang paling rendah. Rendahnya performa pada kelas netral dipengaruhi oleh ketidakseimbangan jumlah data, di mana jumlah data sentimen netral paling sedikit dibandingkan kelas positif dan negatif. Kondisi tersebut menyebabkan model kurang optimal dalam mempelajari pola pada kelas netral. Kasus ketidakseimbangan data ini telah banyak terjadi, bahkan telah dirangkum dalam *systematic literature review* yang dilakukan oleh [42], yang menemukan bahwa masalah ini dapat mempengaruhi akurasi dan performa model klasifikasi. Selain itu, hasil pengujian menunjukkan bahwa pembagian data 90:10 menghasilkan performa terbaik dibandingkan skenario 70:30 dan 80:20 pada seluruh *kernel*. Hal ini menunjukkan bahwa penggunaan jumlah data latih yang lebih besar memberikan kemampuan yang lebih baik bagi model dalam mempelajari pola sentimen. Semakin banyak data latih yang digunakan, maka model memiliki informasi yang lebih memadai untuk membentuk *hyperplane* optimal dalam proses klasifikasi.

Berdasarkan analisis distribusi sentimen, sentimen negatif mendominasi baik pada periode sebelum maupun setelah peresmian Danantara. Dominasi sentimen negatif ini menunjukkan bahwa opini publik pada *platform* X cenderung berisi tanggapan kritis, keraguan, kekhawatiran dan kekecewaan terhadap Danantara. Meskipun demikian, terdapat juga sentimen positif yang menunjukkan adanya dukungan masyarakat terhadap potensi Danantara dalam meningkatkan investasi dan pertumbuhan ekonomi nasional. Temuan mengenai dominannya sentimen negatif ini, didukung oleh performa model klasifikasi yang telah dievaluasi sebelumnya, di mana nilai presisi, *recall*, dan f1-*score* pada kelas negatif menunjukkan hasil relatif tinggi dan konsisten dibandingkan kelas positif dan netral, sehingga model dinilai mampu mengidentifikasi sentimen negatif dengan baik. Dengan demikian, sentimen negatif yang diperoleh dapat dianggap cukup merepresentasikan kecenderungan opini publik.

Penelitian ini memiliki keterbatasan yang mempengaruhi hasil klasifikasi yaitu belum menerapkan teknik *balancing* data, yang menyebabkan performa klasifikasi kelas minoritas masih belum optimal. Oleh karena itu, penelitian selanjutnya disarankan untuk menerapkan metode *balancing* data seperti SMOTE guna meningkatkan keseimbangan distribusi data pada setiap kelas

sentimen. Selain itu, penggunaan pendekatan *aspect-based sentimen analysis* juga dapat dipertimbangkan agar analisis sentimen yang dihasilkan mampu mengidentifikasi aspek-aspek spesifik yang mempengaruhi opini publik terhadap Danantara.

KESIMPULAN DAN REKOMENDASI

Berdasarkan hasil pengujian, didapatkan bahwa metode *Support Vector Machine* (SVM) mampu melakukan klasifikasi sentimen publik terhadap pembentukan Danantara dengan baik melalui perbandingan empat jenis *kernel*. *Kernel polynomial* menghasilkan akurasi tertinggi pada semua skenario pembagian data, terutama pada skenario 90:10 dengan akurasi sebesar 91.66%. selain itu, hasil analisis menunjukkan bahwa sentimen negatif mendominasi opini publik baik sebelum maupun setelah peresmian Danantara.

Penelitian ini memberikan kontribusi dalam pengembangan analisis sentimen, khususnya melalui perbandingan performa beberapa *kernel* SVM pada isu pembentukan Danantara serta analisis distribusi sentimen berdasarkan dua periode waktu yang berbeda. Hasil penelitian menunjukkan bahwa metode SVM memiliki kemampuan yang baik dalam melakukan klasifikasi data teks pada media sosial dengan jumlah data besar. Selain itu, hasil penelitian ini juga dapat dimanfaatkan sebagai bahan evaluasi oleh pemerintah dalam memahami persepsi dan *respon* masyarakat terhadap pembentukan Danantara.

Meskipun demikian, penelitian ini masih memiliki keterbatasan yaitu belum mampu menangani ketidakseimbangan jumlah data antar kelas sentimen. Oleh karena itu, penelitian selanjutnya disarankan untuk menerapkan teknik *balancing* data serta pendekatan *aspect-based analysis* agar mampu menghasilkan analisis sentimen yang lebih mendalam dan akurat.

REFERENSI

[1] M. Ratnayani and D. R. Riadi, "ANALISIS REAKSI PASAR SAHAM BUMN TERHADAP PERESMIAN DANANTARA," *Prosiding Seminar Sosial Politik, Bisnis, Akuntansi dan Teknik*, vol. 7, pp. 187–199, Oct. 2025, doi: 10.32897/sobat.2025.7.1.5246.

[2] H. T. Lumban Gaol, M. Fadil, A. Fahrieza, A. Fahrezi, and S. Azima, "Analisis Hukum Peran dan Kedudukan Daya Anagata Nusantara (Danantara) Sebagai Badan Sovereign Wealth Fund (SWF) Dalam Mengelola Investasi Masa Depan Bangsa," *Recht Studiosum Law Review*, vol. 4, no. 1, pp. 24–36, May 2025, doi: 10.32734/rslr.v4i1.20530.

[3] E. R. Puspapertiwi and I. S. Adhi, "Apa Dampak Langsung Danantara bagi Masyarakat Umum?," *Kompas.com*, Feb. 25, 2025. Accessed: May 29, 2026. [Online]. Available: https://www.kompas.com/tren/read/2025/02/25/050000865/apa-dampak-langsung-danantara-bagi-masyarakat-umum-?page=all&_gl=1*1lndi5j*_ga*YW1wLWVyOGwwcTdXMEVLUGRZTnNHMWY5RGlub%20jE4N0cwR2haT2ZnZkNkam1Ud0VHdy05RUNDeDZJckVFTEx3alpCemY.*_ga_77DJNQ0%20227*MTc0MTY2MTY4Ny4xLjEuMTc0MTY2MTY4OC4wLjAuMA..

[4] D. Arsyanda, S. Rodiah, and A. S. Rohman, "Peran akun autobase X (twitter) dalam memenuhi kebutuhan informasi followers," *Pustaka Karya : Jurnal Ilmiah Ilmu Perpustakaan dan Informasi*, vol. 13, no. 1, pp. 233–241, Jun. 2025, doi: 10.18592/pk.v13i1.15927.

[5] B. Febrianto, M. R. Handayani, N. C. H. Wibowo, and K. Umam, "OPINI PUBLIK DI MEDIA X TERHADAP PATRICK KLUIVERT SEBAGAI PELATIH TIMNAS INDONESIA YANG BARU DENGAN METODE NAÏVE BAYES," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 3, pp. 2571–2583, Sep. 2025, doi: 10.29100/jipi.v10i3.8145.

[6] A. A. Abdillah, "SUPPORT VECTOR MACHINES UNTUK MENYELESAIKAN MASALAH KLASIFIKASI PADA PENGENALAN POLA," *Jurnal Poli-Teknologi*, vol. 18, no. 2, Jul. 2019, doi: 10.32722/pt.v18i2.1432.

[7] M. K. Delimayanti, R. Sari, M. Laya, M. R. Faisal, and P. Pahrul, "Pemanfaatan Metode Multiclass-SVM pada Model Klasifikasi Pesan Bencana Banjir di Twitter," *Edu Komputika Journal*, vol. 8, no. 1, pp. 39–47, Jun. 2021, doi: 10.15294/edukomputika.v8i1.47858.

[8] N. G. A. Dasriani, L. A. G. Pariandi, and I. M. Y. Dharma, "Analisis Sentimen Program Jaminan Kesehatan Nasional Menggunakan Multiclass Support Vector Machine," *BIOS : Jurnal Teknologi Informasi dan Rekayasa Komputer*, vol. 6, no. 1, pp. 20–30, Nov. 2024, doi: 10.37148/bios.v6i1.136.

[9] T. M. Permata Aulia, N. Arifin, and R. Mayasari, "PERBANDINGAN KERNEL SUPPORT VECTOR MACHINE (SVM) DALAM PENERAPAN ANALISIS SENTIMEN VAKSINISASI COVID-19," *SINTECH (Science and Information Technology) Journal*, vol. 4, no. 2, pp. 139–145, Oct. 2021, doi: 10.31598/sintechjournal.v4i2.762.

[10] S. Styawati, N. Hendrastuty, and A. R. Isnain, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 6, no. 3, pp. 150–155, Oct. 2021, doi: 10.30591/jpit.v6i3.2870.

[11] J. Anggraini and D. Alita, "Implementasi Metode SVM Pada Sentimen Analisis Terhadap Pemilihan Presiden (Pilpres) 2024 Di Twitter," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 9, no. 2, pp. 102–111, Aug. 2024, doi: 10.30591/jpit.v9i2.6560.

[12] Y. Ansori and K. F. H. Holle, "Perbandingan Metode Machine Learning dalam Analisis Sentimen Twitter," *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 10, no. 4, p. 429, Dec. 2022, doi: 10.26418/justin.v10i4.51784.

[13] J. E. Br Sinulingga and H. C. K. Sitorus, "Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 14, no. 1, pp. 42–53, Feb. 2024, doi: 10.34010/jamika.v14i1.11946.

[14] A. F. Njurumana, F. Hariadi, and N. B. Uly, "Analisis Sentimen Pengguna X Terkait Danantara Menggunakan Metode Support Vector Machine," *Riau Jurnal Teknik Informatika*, vol. 4, no. 2, Jul. 2025, doi: 10.30606/rjti.v4i2.3439.

[15] S. A. Nugraha, "PENERAPAN LEXICON BASED UNTUK ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP DANANTARA," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 3, pp. 4949–4957, May 2025, doi: 10.36040/jati.v9i3.13836.

[16] M. L. Hermanto, F. Fathoni, O. Ardhillah, and A. Ibrahim, "ANALISIS SENTIMEN MASYARAKAT TERHADAP DANANTARA DI PLATFORM X DENGAN METODE SVM," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 4, pp. 6779–6785, May 2025, doi: 10.36040/jati.v9i4.14189.

[17] F. P. Agustinus and M. A. D. Ardiansyah, "Analisis Sentimen Opini Publik terhadap BPI Danantara di Media Sosial X Menggunakan IndoRoBERTa dan SVM," *Seminar Nasional Teknologi & Sains*, vol. 5, no. 1, pp. 235–242, Jan. 2026.

[18] M. E. Siregar, S. Dermawan, and A. A. Hisyam, "PERBANDINGAN KINERJA NAIVE BAYES DAN KNN DALAM ANALISIS SENTIMEN KOMENTAR X DENGAN DAN TANPA TEXT PREPROCESSING (STUDI KASUS: DANANTARA)," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3, Jul. 2025, doi: 10.23960/jitet.v13i3.6612.

[19] Z. A. Fernando, Rumini, and T. Susanto, "Public Sentiment Analysis on the Launch of Danantara Using Support Vector Machine (SVM) Algorithm," *Journal of Artificial Intelligence and Software Engineering*, vol. 5, no. 4, pp. 1425–1433, Dec. 2025.

[20] Welda, A. S. Kusuma, and I. P. R. L. Putra, "Sentiment Analysis of Public Opinion on Danantara Via Social Media X," *Jurnal Krisnadana (Jurnal Komputer, Sistem Kendali & Jaringan)*, vol. 5, no. 3, pp. 598–605, May 2026.

[21] A. Y. Pratama, G. A. Sanjaya, N. K. Lubis, M. R. Aditya, and Yennimar, "Analisis Sentimen Publik Terkait Danantara Menggunakan Algoritma IndoBERT pada Platform Media Sosial," *METIK JURNAL*, vol. 9, no. 1, pp. 92–100, 2025.

[22] M. Fazri and A. Voutama, "ANALISIS SENTIMEN PUBLIK TERHADAP DANANTARA DI MEDIA SOSIAL X MENGGUNAKAN NLP DAN PEMBELAJARAN MESIN," *JOISIE (Journal Of Information Systems And Informatics Engineering)*, vol. 9, no. 1, p. 197, Jul. 2025, doi: 10.35145/joisie.v9i1.4924.

[23] S. R. Hakim, M. A. Rizki, N. I. Zekha F, N. Fitri, Y. R. A, and R. Nooraeni, "ANALISIS SENTIMEN PENGGUNA INSTAGRAM TERHADAP KEBIJAKAN KEMDIKBUD MENGENAI BANTUAN KUOTA INTERNET DENGAN METODE SUPPORT VECTOR MACHINE (SVM)," *Jurnal MSA (Matematika dan Statistika serta Aplikasinya)*, vol. 8, no. 2, p. 15, Dec. 2020, doi: 10.24252/msa.v8i2.16795.

[24] S. Awaliah, A. N. Ramadini, N. F. Zetti, A. Septiarini, and N. Puspitasari, "Perbandingan Metode Naïve Bayes dan Support Vector Machine dalam Analisis Sentimen Citra Polri berdasarkan Opini pada Platform Twitter/X," *Jurnal Sains Teknologi dan Sistem Informasi*, vol. 6, no. 1, pp. 52–59, Apr. 2026.

[25] D. Pratiwi, N. Khoerani, and S. Sari, "Analisis Sentimen Terhadap Kebijakan SubsidiPembelian Kendaraan Bertenaga Listrik Di Indonesia Menggunakan Pendekatan Inset Lexicon dan Metode Support Vector Machine," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 6, pp. 1303–1314, Dec. 2025.

[26] M. F. N. Fathoni, E. Y. Puspaningrum, and A. N. Sihananto, "Perbandingan Performa Labeling Lexicon InSet dan VADER pada Analisa Sentimen Rohingya di Aplikasi X dengan SVM," *Modem : Jurnal Informatika dan Sains Teknologi.*, vol. 2, no. 3, pp. 62–76, Jul. 2024, doi: 10.62951/modem.v2i3.112.

[27] F. Koto and G. Y. Rahmaningtyas, "Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs," in *2017 International Conference on Asian Language Processing (IALP)*, IEEE, Dec. 2017, pp. 391–394. doi: 10.1109/IALP.2017.8300625.

[28] H. Firda, P. Putra, N. R. Oktadini, P. E. Sevtiyuni, and A. Meiriza, "Comparison of Rating-based and Inset Lexicon-based Labeling in Sentiment Analysis using SVM (Case Study: GoBiz Application Reviews on Google Play Store)," *SISTEMASI*, vol. 14, no. 2, p. 516, Mar. 2025, doi: 10.32520/stmsi.v14i2.4795.

[29] M. F. N. Fathoni, E. Y. Puspaningrum, and A. N. Sihananto, "Perbandingan Performa Labeling Lexicon InSet dan VADER pada Analisa Sentimen Rohingya di Aplikasi X dengan SVM," *Modem : Jurnal Informatika dan Sains Teknologi.*, vol. 2, no. 3, pp. 62–76, Jul. 2024, doi: 10.62951/modem.v2i3.112.

[30] D. Pratiwi, N. Khoerani, and S. Sari, "Analisis Sentimen Terhadap Kebijakan Subsidi Pembelian Kendaraan Bertenaga Listrik Di Indonesia Menggunakan Pendekatan Inset Lexicon dan Metode Support Vector Machine," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 6, pp. 1303–1314, Dec. 2025.

[31] M. S. Al Markas, A. Siska, and L. B. Ilmuwan, "Implementasi Fitur Vector Bag Of Word Dan TF IDF untuk Analisis Sentiment," *LINIER: Literatur Informatika dan Komputer*, vol. 2, no. 2, pp. 136–146, Sep. 2025, doi: 10.33096/linier.v2i2.3104.

[32] C. E. Amos Pah and D. N. Utama, "Decision Support Model for Employee Recruitment Using Data Mining Classification," *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 5, pp. 1511–1516, May 2020, doi: 10.30534/ijeter/2020/06852020.

[33] L. Hernando, V. A. Handayani, D. P. Caniago, and N. W. Nasution, "Penerapan data mining dalam analisa profil mahasiswa menggunakan metode support vector machines (SVM)," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 4, no. 2, pp. 477–483, Sep. 2023, doi: 10.37859/coscitech.v4i2.5107.

[34] A. Putra G, M. A. Tiro, and M. K. Aidid, "Metode Boostrap dan Jackknife dalam Mengestimasi Parameter Regresi Linear Ganda (Kasus: Data Kemiskinan Kota Makassar Tahun 2017),"

VARIANSI: Journal of Statistics and Its application on Teaching and Research, vol. 1, no. 2, p. 32, Jul. 2019, doi: 10.35580/variansiunm12895.

[35] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, "Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 153–160, Oct. 2023, doi: 10.57152/malcom.v3i2.897.

[36] N. Kamilatutsaniya, E. Daniati, and M. N. Muzaki, "Pemodelan Klasifikasi Popularitas Produk Skincare Menggunakan Support Vector Machine (SVM): Studi Komparatif Kinerja Kernel.," *The Indonesian Journal of Computer Science Research*, vol. 4, no. 2, pp. 86–95, Jul. 2025, doi: 10.59095/ijcsr.v4i2.209.

[37] Suyanto, *Data Mining Untuk Klasifikasi Dan Klasterisasi Data Edisi Revisi*. Bandung: Informatika Bandung, 2019.

[38] N. Nursya'bani, Ruliana, and M. K. Aidid, "APPLICATION OF SVM FOR SENTIMENT ANALYSIS REGARDING THE EFFICIENCY OF APBN AND APBD IN 2025," *Journal of Statistics and Its Application on Teaching and Research*, vol. 8, no. 1, pp. 83–95, Apr. 2026.

[39] Badriyah, T. Chamidy, and Suhartono, "Application of SMOTE in Sentiment Analysis of MyXL User Reviews on Google Play Store," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 10, no. 1, pp. 74–86, Jan. 2025, doi: 10.14421/jiska.2025.10.1.74-86.

[40] I. Frianda Gultom, Sriani, and Suhardi, "Analisis Sentimen Kebijakan Pemberian Subsidi Motor Listrik Menggunakan Metode Support Vector Machine," *JURNAL FASILKOM*, vol. 13, no. 3, pp. 511–517, Dec. 2023, doi: 10.37859/jf.v13i3.6225.

[41] A. F. Abdul Fadlil, I. Riadi, and F. Andrianto, "Improving Sentiment Analysis in Digital Marketplaces through SVM Kernel Fine-Tuning," *International Journal of Computing and Digital Systems*, vol. 15, no. 1, pp. 159–171, Jul. 2024, doi: 10.12785/ijcds/160113.

[42] Rifqi Fitriadi and Deni Mahdiana, "SYSTEMATIC LITERATURE REVIEW OF THE CLASS IMBALANCE CHALLENGES IN MACHINE LEARNING," *Jurnal Teknik Informatika (Jutif)*, vol. 4, no. 5, pp. 1099–1107, Oct. 2023, doi: 10.52436/1.jutif.2023.4.5.970.