

PENERAPAN LEXICON BASED UNTUK ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP DANANTARA

Satria Adi Nugraha

Sistem Informasi, Universitas Ciputra Surabaya
CitraLand CBD Boulevard Surabaya, Indonesia
satria.nugraha@ciputra.ac.id

ABSTRAK

Investasi strategis melalui sovereign wealth fund (SWF) menjadi kebijakan yang semakin banyak diterapkan oleh berbagai negara, termasuk Indonesia yang baru saja meresmikan Badan Pengelola Investasi Daya Anagata Nusantara (Danantara) pada Februari 2025. Meskipun memiliki tujuan utama untuk meningkatkan efisiensi pengelolaan aset strategis dan menarik investasi global, keberadaan Danantara menimbulkan beragam opini publik, terutama di media sosial seperti Twitter. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat Indonesia terhadap Danantara menggunakan metode *Lexicon-Based Sentiment Analysis*. Penelitian ini diharapkan dapat memberikan wawasan bagi pemerintah dan pemangku kepentingan dalam meningkatkan strategi komunikasi, transparansi kebijakan, serta membangun kepercayaan publik terhadap pengelolaan investasi nasional. Data dikumpulkan dari Twitter dan diolah melalui tahapan *pre-processing* serta pencocokan dengan kamus sentimen berbahasa Indonesia. Hasil analisis menunjukkan bahwa sentimen publik terhadap Danantara cenderung beragam dengan dominasi negatif (44.9%) sedikit lebih banyak daripada positif (44.1%) dan sisanya merupakan sentimen netral (14%). Faktor-faktor seperti transparansi kebijakan dan persepsi terhadap pengelolaan investasi berkontribusi terhadap variasi sentimen ini.

Kata kunci : *Sovereign Wealth Fund*, Sentimen Publik, Media Sosial, Danantara

1. PENDAHULUAN

Dalam upaya memperkuat perekonomian nasional, Indonesia memperkenalkan Badan Pengelola Investasi Daya Anagata Nusantara (Danantara), sebuah *sovereign wealth fund* (SWF) yang berperan dalam mengelola dan mengoptimalkan investasi negara. Diresmikan pada 24 Februari 2025, Danantara bertujuan untuk meningkatkan efisiensi pengelolaan aset strategis nasional serta menarik investasi global guna mempercepat pembangunan ekonomi [1][2].

Sebagai kebijakan strategis, pembentukan Danantara tidak luput dari perhatian publik. Kebijakan pemerintah yang berdampak luas terhadap sektor ekonomi dan sosial sering kali memunculkan berbagai reaksi di masyarakat, mulai dari dukungan hingga kritik. Seiring dengan meningkatnya penggunaan media sosial, *platform* seperti Twitter telah menjadi salah satu kanal utama bagi masyarakat dalam menyampaikan opini dan persepsi terhadap kebijakan pemerintah [3].

Analisis sentimen telah menjadi pendekatan yang banyak digunakan untuk memahami persepsi publik terhadap suatu isu. Salah satu metode analisis sentimen adalah menggunakan Lexicon Based yang memungkinkan klasifikasi sentimen berdasarkan kata dan ekspresi yang digunakan dalam teks dibandingkan kamus yang sudah ada. Kamus tersebut terdiri dari kamus kata-kata positif dan negatif. Dengan melakukan pengolahan data yang sesuai, metode ini telah terbukti efektif dalam menangkap nuansa sentimen dalam bahasa alami, terutama dalam konteks analisis opini di media sosial[4][5][6].

Dalam perkembangan analisis sentimen, pendekatan *lexicon-based* menunjukkan keunggulan

signifikan dibandingkan metode lain seperti *machine learning* dan *large language models* (LLM), terutama dalam hal generalisasi lintas domain, interpretabilitas, dan efisiensi sumber daya. Pendekatan lexicon tidak hanya menunjukkan performa akurat (akurasi >75%) tetapi juga memberikan kemampuan untuk menangkap gradasi sentimen yang lebih halus, yang tidak tersedia pada model biner berbasis ML maupun LLM [7][8].

Dengan demikian, pendekatan lexicon-based tidak hanya unggul dari sisi akurasi dalam konteks tertentu, tetapi juga menawarkan nilai tambah berupa interpretasi yang transparan dan dapat diandalkan bagi keperluan ilmiah dan praktis lintas domain.

Penelitian ini bertujuan untuk menganalisis sentimen masyarakat Indonesia terhadap Danantara dengan menggunakan metode *lexicon-based* pada data Twitter berbahasa Indonesia. Dengan memahami sentimen publik, penelitian ini diharapkan dapat memberikan wawasan bagi pengambil kebijakan dan pengelola Danantara dalam menyusun strategi komunikasi, meningkatkan transparansi, serta membangun kepercayaan publik terhadap kebijakan investasi negara.

2. TINJAUAN PUSTAKA

2.1. Danantara

Danantara merupakan sebuah *Sovereign Wealth Fund* (SWF). SWF merupakan instrumen investasi yang dikelola oleh negara untuk mengelola kekayaan nasional dan mengoptimalkan aset negara dalam jangka panjang. SWF sering kali dibentuk dengan tujuan untuk mendiversifikasi sumber pendapatan negara, mengelola surplus fiskal, serta memperkuat daya saing ekonomi di tingkat global [9].

Danantara dibentuk berdasarkan Peraturan Pemerintah Republik Indonesia Nomor 10 tahun 2025. Danantara mengonsolidasikan sejumlah perusahaan BUMN strategis, termasuk Bank Mandiri, BRI, PLN, Pertamina, BNI, Telkom Indonesia, dan MIND ID. Total aset yang dikelola diproyeksikan mencapai Rp14,72 kuadriliun (US\$900 miliar), menjadikannya salah satu instrumen investasi terbesar di kawasan Asia Tenggara [1].

Pembentukan Danantara dipandang sebagai langkah penting dalam menghadapi tantangan ekonomi global, terutama dalam meningkatkan daya saing Indonesia di tengah ketidakpastian pasar dunia. Optimalisasi aset negara melalui SWF dapat memberikan dampak positif dalam memperkuat stabilitas fiskal, menarik investasi asing, serta mendukung agenda pembangunan jangka panjang [10][11]. Namun, keberhasilan Danantara sangat bergantung pada transparansi tata kelola, mitigasi risiko investasi, serta tingkat kepercayaan publik terhadap kebijakan yang diimplementasikan.

2.2. Lexicon Based Sentiment Analysis

Analisis sentimen berbasis leksikon adalah pendekatan yang memanfaatkan kamus sentimen untuk menilai polaritas emosional dalam teks. Metode ini bergantung pada daftar kata yang telah diberi bobot sentimen tertentu, memungkinkan identifikasi opini positif, negatif, atau netral dalam sebuah dokumen atau kalimat. Keunggulan utama pendekatan ini adalah kemudahannya dalam interpretasi dan implementasi, karena tidak memerlukan data *training* yang besar dibanding metode *machine learning* lainnya seperti Support Vector Machine (SVM) dan Naïve Bayes [7].

Proses analisis sentimen berbasis leksikon umumnya melibatkan beberapa tahap. Pertama, teks yang akan dianalisis melalui tahap *preprocessing*, seperti penghapusan tanda baca, normalisasi kata, dan *tokenization*. Selanjutnya, setiap kata dalam teks dicocokkan dengan kamus sentimen untuk memperoleh skor sentimen yang sesuai. Skor-skor ini kemudian dijumlahkan atau dirata-rata untuk menentukan polaritas keseluruhan teks. Pendekatan ini efektif dalam menangani teks informal, seperti postingan media sosial, di mana penggunaan bahasa cenderung lebih variatif [12].

Kamus sentimen berisi daftar kata yang masing-masing diberi bobot yang mencerminkan intensitas emosionalnya. Bobot ini pada umumnya berupa nilai numerik yang menunjukkan seberapa positif atau negatif suatu kata. Sebagai contoh, kata dengan konotasi sangat positif mungkin diberi skor +5, sedangkan kata dengan konotasi sangat negatif diberi skor -5. Penentuan bobot ini dapat dilakukan secara manual oleh pakar bahasa atau melalui pendekatan statistik yang menganalisis frekuensi dan konteks penggunaan kata dalam korpus teks yang besar.

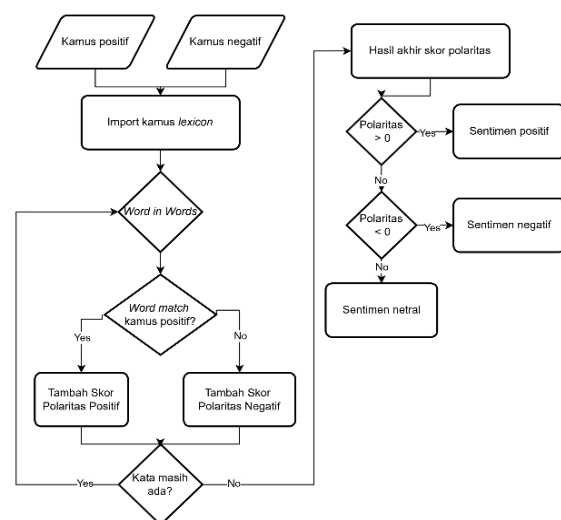
Proses pelabelan dilakukan menggunakan pendekatan *lexicon based*, di mana hasil klasifikasi dianalisis untuk mengidentifikasi jenis opini yang

paling dominan. Setiap kalimat dianalisis untuk mengekstraksi nilai sentimen, kemudian kata-kata dalam teks diklasifikasikan menggunakan persamaan (1).

$$S = \sum_{i=1}^n w_i \quad (1)$$

$$\begin{aligned} S_{positive} &= S > 0 \\ S_{negative} &= S < 0 \\ S_{neutral} &= S = 0 \end{aligned} \quad (2)$$

Dengan demikian, total skor positif, negatif, dan netral dapat ditentukan berdasarkan kata-kata yang membentuk teks. Persamaan (2) digunakan untuk mengidentifikasi orientasi sentimen dengan membandingkan proporsi nilai positif, negatif, dan netral.



Gambar 1. Flowchart Lexicon Based

Gambar *flowchart* pada Gambar 1 menggambarkan mekanisme dasar dari pendekatan *lexicon-based* dalam analisis sentimen, yang banyak digunakan dalam studi-studi terdahulu karena sifatnya yang transparan, interpretatif, dan hemat sumber daya. Proses diawali dengan pemuatan dua kamus—positif dan negatif—yang berisi daftar kata beserta skor polaritasnya masing-masing. Setiap kata dalam teks yang dianalisis kemudian dicocokkan dengan kedua kamus. Jika kata ditemukan dalam kamus positif, maka skor polaritas positif ditambahkan; sebaliknya, jika terdapat dalam kamus negatif, maka skor polaritas negatif dikurangi. Proses ini berulang hingga seluruh kata selesai dievaluasi. Hasil akhir berupa akumulasi skor polaritas, yang kemudian diklasifikasikan menjadi sentimen positif jika nilainya lebih dari nol, negatif jika kurang dari nol, dan netral jika sama dengan nol. Pendekatan ini banyak digunakan dalam penelitian terdahulu karena kesederhanaannya dan kemampuannya untuk menyesuaikan dengan konteks lokal melalui pengembangan kamus sendiri.

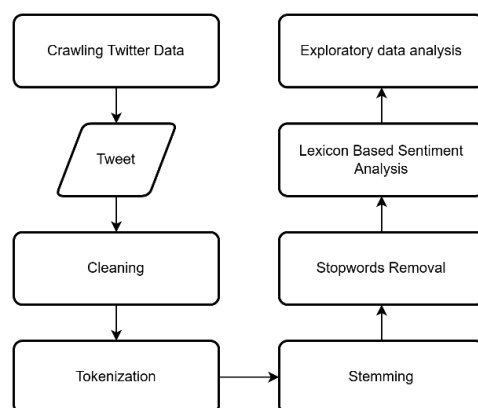
2.3. Penelitian Terdahulu

Analisis sentimen terhadap kebijakan pemerintah Indonesia telah menjadi topik penelitian yang signifikan dalam beberapa tahun terakhir. Penelitian-penelitian sebelumnya umumnya menggunakan metode machine learning untuk mengevaluasi opini publik. Machine learning dengan beberapa algoritma seperti Linear SVC, SVM, dan Decision Tree digunakan untuk mengevaluasi sentiment dari *feedback* masyarakat terhadap peraturan pemerintah di Twitter dengan akurasi yang cukup baik menunjukkan efektivitas beberapa metode tersebut dalam klasifikasi sentimen [3]. Support Vector Machine (SVM) dengan metode hybrid juga dilakukan untuk meneliti *feedback* dan *review* pelanggan terbukti akurat untuk melakukan klasifikasi sentimen [13].

Meskipun metode machine learning telah banyak digunakan, pendekatan *lexicon based* menawarkan keunggulan dalam interpretabilitas dan kemudahan implementasi. Penelitian menggunakan *lexicon based* untuk analisis sentimen terkait Islamophobia di media sosial, menunjukkan bahwa efektif dalam menangkap nuansa sentimen dalam teks informal [14]. *Lexicon-based* juga terbukti akurat untuk melakukan analisa sentimen TikTok *review* yang digabungkan dengan model SVM [5] serta sentimen publik di Twitter terkait isu COVID-19 selama pandemi yang hasilnya dapat membantu pemerintah dalam merumuskan kebijakan yang lebih responsif [15].

Penerapan *lexicon based* dalam konteks analisis sentimen terhadap kebijakan pemerintah Indonesia terlihat masih terbatas. Selain itu, belum ada penelitian yang secara spesifik mengevaluasi sentimen publik terhadap Badan Pengelola Investasi Daya Anagata Nusantara (Danantara) menggunakan metode *lexicon based*. Oleh karena itu, terdapat celah penelitian yang signifikan untuk mengaplikasikan metode *lexicon based* dalam menganalisis sentimen masyarakat Indonesia terhadap Danantara, yang dapat memberikan wawasan baru mengenai persepsi publik terhadap kebijakan investasi nasional ini dengan lebih cepat.

3. METODE PENELITIAN



Gambar 2. Alur Penelitian

Alur penelitian pada penelitian ini digambarkan pada gambar 2 di atas. Penelitian diawali dengan melakukan *crawling* data dari Twitter berupa komentar, jumlah *quote*, jumlah *reply*, dan jumlah *retweet*. Tweet yang diambil adalah data *tweet* menggunakan Bahasa Indonesia yang mengandung kata kunci “Danantara” dari periode 23 Februari sampai 7 Maret 2025. Hal ini mencerminkan opini langsung masyarakat karena bertepatan dengan H-1 peresmian Danantara dan 11 hari pasca peresmian Danantara. Data yang berhasil dikumpulkan adalah sebanyak 5337 data.

Tabel 1. Deskripsi Data

Nama Kolom	Deskripsi
Created_at	Hari, tanggal, dan jam <i>tweet</i> tersebut di posting
Full_text	Isi <i>tweet</i> atau komentar
Quote_count	Jumlah <i>tweet</i> tersebut di <i>quote</i>
reply_count	Jumlah <i>reply</i> pada <i>tweet</i> tersebut
retweet_count	Jumlah <i>retweet</i> pada <i>tweet</i> tersebut

Tabel 1 merupakan atribut dari data yang diambil. Mulai dari *created_at* yang berisi keterangan hari, tanggal, dan jam saat komentar tersebut diunggah sampai ke *retweet_count* yang berisi jumlah *retweet* pada komentar tersebut. Atribut-atribut selain *full_text* akan diolah pada bagian *exploratory data analysis* untuk melihat hubungan hasil sentimen dengan faktor-faktor lainnya.

Data yang sudah dikumpulkan kemudian di olah melalui beberapa tahap. Tahap pertama merupakan *cleaning data*. Tahap ini membersihkan data atau komentar dari bagian-bagian teks yang tidak terpakai, seperti menghilangkan *account tag*, *hyperlink*, angka, spasi berlebihan, dan *remove punctuation*.

Data komentar yang sudah bersih kemudian dilakukan tokenisasi. Tahap tokenisasi merupakan proses memisahkan kalimat menjadi beberapa bagian token [16]. Token pada penelitian ini menggunakan format kata-kata yang dipecah dari sebuah kalimat. Tahap ini penting untuk memudahkan tahap selanjutnya yaitu *stemming* dan *stopwords removal*.

Stemming merupakan metode yang secara proses morfologi mengurangi sebuah kata ke bentuk paling standar atau awal dengan cara mengurangi imbuhan baik di depan, belakang, maupun di tengah dari sebuah kata. Penggunaan *stemming* penting dilakukan dalam proses analisis sentimen karena proses ini mengurangi ukuran dari kamus yang dipakai. Hal itu secara langsung mengurangi biaya komputasi dan meningkatkan performa system yang dipakai sehingga analisis yang dilakukan lebih sesuai [17].

Stopwords removal adalah merupakan tahap penghapusan kata yang tidak informatif atau tidak punya makna dalam proses *text classification* yang dijalankan pada saat *preprocessing*. *Stopwords* sendiri biasanya sering muncul dalam banyak bahasa dalam sebuah dokumen ataupun data yang sebenarnya hanya membawa sedikit makna mengenai bagian dari

dokumen atau teks tersebut [18]. Beberapa *stopwords* dalam Bahasa Indonesia antara lain seperti kata 'yang', 'untuk', 'lagi', 'maka', 'dan', 'kami', dan masih banyak kata lainnya.

Data yang sudah melewati tahap *preprocessing* mulai dari *cleaning*, *tokenization*, *stemming*, dan *stopwords removal* artinya sudah siap untuk dilakukan klasifikasi teks yaitu untuk analisis sentimen. Proses pelabelan analisis sentimen pada penelitian ini menggunakan *lexicon based* yang artinya mencocokkan teks dengan kamus yang sudah ada.

Penelitian ini menggunakan InSet (*Indonesian Sentiment Lexicon*). InSet terdiri dari 3.609 kata positif dan 6.609 kata negatif, masing-masing dengan bobot yang berkisar antara -5 hingga +5. Kamus ini dibangun dengan mengumpulkan kata-kata dari tweet berbahasa Indonesia, yang kemudian diberi bobot secara manual dan diperluas dengan menambahkan stemming serta sinonim. Penggunaan InSet mampu meningkatkan akurasi analisis sentimen pada mikroblog berbahasa Indonesia [19]. InSet ini juga cocok dengan penelitian ini karena sama-sama menggunakan kata-kata dari *tweet* Bahasa Indonesia sehingga memungkinkan makna yang sama.

Hasil dari analisis sentimen menggunakan *lexicon based* ini akan keluar label sentimen positif, netral, dan negatif. Label-label tersebut kemudian dianalisa lebih lanjut dihubungkan dengan atribut lain seperti *quote_count*, *reply_count*, *retweet_count* dan *created_at*. Hubungan tersebut divisualisasikan sehingga bisa mendapatkan informasi tambahan terkait sentimen Danantara ini.

4. HASIL DAN PEMBAHASAN

Hasil pada penelitian ini dijabarkan dalam sebuah 2 bagian. Pada bagian pertama akan menjelaskan hasil analisis sentimen dari *tweet* mengenai Danantara menggunakan *lexicon based* dari InSet. Bagian kedua akan menjelaskan visualisasi hubungan antara hasil sentimen dengan atribut dari sebuah *tweet*.

4.1. Klasifikasi Sentimen Danantara

Bagian ini menunjukkan hasil dari analisa sentimen dari *tweet* Bahasa Indonesia mengenai Danantara menggunakan metode *lexicon* memanfaatkan kamus InSet.

Gambar 3 memperlihatkan implementasi proses analisis sentimen berbasis kamus (*lexicon-based*) dalam bahasa pemrograman Python menggunakan *library* pandas. Tahapan pertama dalam metode ini adalah memuat dua file kamus sentimen, yaitu *kamus_positive.xlsx* dan *kamus_negative.xlsx*, yang masing-masing berisi kata-kata dengan skor polaritas positif dan negatif. File tersebut dibaca dan dikonversi menjadi dua struktur data *dictionary*, yaitu *lexicon_positive_dict* dan *lexicon_negative_dict*, yang berfungsi sebagai acuan dalam penghitungan skor sentimen. Proses ini bersifat manual dan berbasis pada

daftar kata yang telah disiapkan berdasarkan kamus InSet.

LEXICON BASED - INSET

```
lexicon_positive = pd.read_excel('/content/drive/My Drive/kamus_positive.xlsx')
lexicon_positive_dict = {}
for index, row in lexicon_positive.iterrows():
    if row[0] not in lexicon_positive_dict:
        lexicon_positive_dict[row[0]] = row[1]

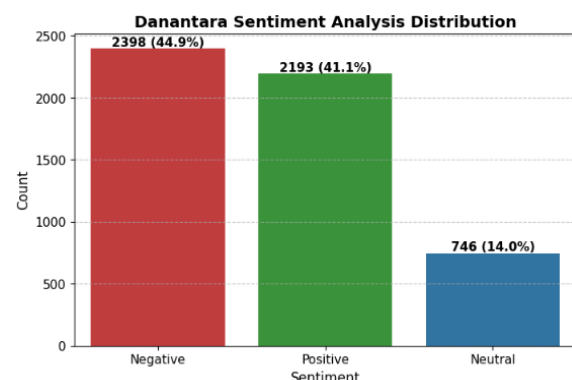
lexicon_negative = pd.read_excel('/content/drive/My Drive/kamus_negative.xlsx')
lexicon_negative_dict = {}
for index, row in lexicon_negative.iterrows():
    if row[0] not in lexicon_negative_dict:
        lexicon_negative_dict[row[0]] = row[1]

def sentiment_analysis_lexicon_indonesia(text):
    score = 0
    for word in text:
        if word in lexicon_positive_dict:
            score = score + lexicon_positive_dict[word]
    for word in text:
        if word in lexicon_negative_dict:
            score = score + lexicon_negative_dict[word]
    sentimen = ''
    if (score > 0):
        sentimen = 'Positive'
    elif (score < 0):
        sentimen = 'Negative'
    else:
        sentimen = 'Neutral'
    return score, sentimen

[ ] results = data['filtered_text'].apply(sentiment_analysis_lexicon_indonesia)
results = list(zip(*results))
data['polarity_score'] = results[0]
data['sentiment'] = results[1]
```

Gambar 3. Tahapan *Lexicon Based* menggunakan Inset

Terdapat fungsi yang digunakan untuk menganalisis sentimen dari setiap teks yang telah melalui proses *pre-processing*. Skor sentimen dihitung dengan menelusuri setiap kata dalam teks dan menjumlahkan nilai-nilai yang sesuai berdasarkan keberadaannya dalam kamus positif atau negatif. Jika skor akhir lebih dari nol, maka teks diklasifikasikan sebagai 'Positive', jika kurang dari nol maka 'Negative', dan jika nol maka 'Neutral'. Hasil dari fungsi ini kemudian diaplikasikan ke seluruh data menggunakan metode *.apply()* dan disimpan dalam kolom baru bernama *polarity_score* dan *sentiment*. Proses ini menunjukkan bahwa metode *lexicon-based* menggunakan InSet dapat menghasilkan output yang transparan dan dapat ditelusuri



Gambar 2. Distribusi Sentimen Danantara

Gambar 2 menjelaskan Hasil analisis sentimen terhadap Danantara. Dari gambar tersebut menunjukkan bahwa mayoritas opini yang terdeteksi dari data yang dikumpulkan bersifat negatif. Dari total

data yang dianalisis, sebanyak 2.398 atau 44,9% dari seluruh opini yang terekam mengandung sentimen negatif. Proporsi ini mengindikasikan bahwa terdapat kekhawatiran atau ketidakpuasan yang cukup dominan dalam diskusi mengenai Danantara di Twitter. Sentimen negatif ini bisa disebabkan oleh berbagai faktor, seperti persepsi publik terhadap layanan atau produk, atau adanya isu tertentu yang menjadi perhatian masyarakat.

Di sisi lain, sentimen positif yang teridentifikasi mencapai 2.193 atau 41,1% dari total data, yang menunjukkan bahwa meskipun opini negatif cukup dominan, masih terdapat proporsi signifikan dari pengguna atau masyarakat yang memberikan respons positif terhadap Danantara. Hal ini mengindikasikan bahwa di tengah kritik yang ada, masih terdapat banyak masyarakat yang memiliki pandangan positif terhadap kebijakan pemerintah mengenai Danantara ini. Perbandingan yang cukup berimbang antara sentimen negatif dan positif menunjukkan adanya persepsi yang beragam dalam masyarakat.

Sementara itu, sentimen netral memiliki jumlah yang lebih kecil dibandingkan dua kategori lainnya, dengan 746 *tweet* atau sekitar 14,0% dari total data. Sentimen netral umumnya mencerminkan opini yang tidak condong pada aspek positif maupun negatif, seperti penyampaian informasi atau berita, diskusi objektif, atau pernyataan yang tidak menunjukkan emosi tertentu. Proporsi sentimen netral yang relatif kecil dapat mengindikasikan bahwa pembicaraan mengenai Danantara cenderung bersifat emosional, dengan banyak pengguna yang mengekspresikan opini mereka dalam bentuk dukungan atau kritik yang jelas.

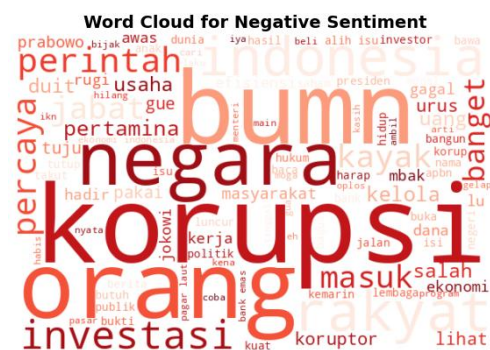


Gambar 3. *Word Cloud* Sentimen Positif Danantara

Word cloud untuk sentimen positif terhadap Danantara seperti yang ditunjukkan pada Gambar 3 menunjukkan bahwa kata-kata yang paling dominan dalam opini masyarakat berkaitan dengan investasi, percaya, dukung, dan ekonomi. Hal ini mengindikasikan bahwa sentimen positif terhadap Danantara banyak dikaitkan dengan aspek kepercayaan dan dukungan terhadap investasi, baik dalam konteks ekonomi nasional maupun peran BUMN. Masyarakat menganggap bahwa hal positif dari Danantara ini adalah adanya investasi yang didukung dengan kepercayaan masyarakat terhadap pemerintah. Kata-kata seperti pemerintah, Indonesia,

dan strategi juga muncul secara signifikan, yang menunjukkan adanya keyakinan bahwa Danantara bisa memiliki kontribusi terhadap pembangunan ekonomi dan kebijakan strategis yang berdampak luas.

Selain itu, kemunculan kata-kata seperti inovasi, maju, bangun, dan manfaat mengisyaratkan bahwa banyak masyarakat yang melihat Danantara sebagai bagian dari kemajuan dan pengembangan ekonomi yang positif. Dukungan terhadap peran Danantara dalam sektor ekonomi tampak dari kehadiran istilah seperti dana, transparansi, dan efisiensi, yang mencerminkan harapan publik terhadap pengelolaan investasi yang baik dan berdampak nyata bagi masyarakat. Secara keseluruhan, analisis ini menunjukkan bahwa sentimen positif terhadap Danantara erat kaitannya dengan optimisme terhadap pertumbuhan ekonomi, kebijakan investasi yang kuat, serta kepercayaan masyarakat terhadap peran pemerintah dalam mewujudkan kesejahteraan bersama.



Gambar 4. *Word Cloud* Sentimen Negatif Danantara

Word cloud untuk sentimen negatif terhadap Danantara ditunjukkan pada Gambar 4. Gambar tersebut menunjukkan bahwa kata korupsi, negara, orang, BUMN, dan investasi menjadi yang paling dominan dalam opini masyarakat. Keberadaan kata korupsi yang sangat menonjol mengindikasikan bahwa ada kekhawatiran besar terhadap transparansi dan tata kelola dalam pengelolaan investasi serta peran BUMN. Selain itu, kata-kata seperti masuk, salah, dan kelola juga mengisyaratkan adanya persepsi negatif terhadap bagaimana proyek atau dana dikelola, yang mungkin dikaitkan dengan isu-isu penyalahgunaan wewenang atau ketidakjelasan kebijakan.

Kemunculan kata seperti pemerintah, Pertamina, politik, dan APBN menunjukkan bahwa banyak opini negatif terhadap Danantara berkaitan dengan kebijakan negara dan bagaimana pemerintah menangani proyek ini. Hal ini dapat mencerminkan ketidakpuasan publik terhadap pengelolaan keuangan negara atau ketidakpercayaan terhadap institusi terkait. Selain itu, kata-kata seperti rugi, usaha, dan ekonomi mengindikasikan adanya kecemasan terkait dampak ekonomi yang dirasakan masyarakat, baik dalam skala individu maupun nasional. Dengan demikian, analisis sentimen negatif ini menyoroti adanya kekhawatiran masyarakat terhadap aspek

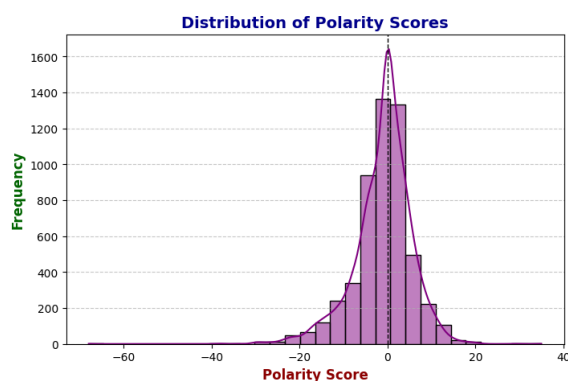
transparansi, tata kelola, dan dampak ekonomi dari Danantara.

Distribusi sentimen ini memberikan gambaran mengenai bagaimana masyarakat menanggapi Danantara di *Twitter*. Dengan tingginya proporsi sentimen negatif, penting bagi pemangku kepentingan terkait untuk mengevaluasi lebih lanjut faktor-faktor yang berkontribusi terhadap persepsi negatif ini. Pendekatan berbasis data dapat digunakan untuk mengidentifikasi pola kritik yang dominan dan merumuskan strategi yang lebih tepat guna dalam meningkatkan citra dan kepercayaan masyarakat.

Secara keseluruhan, meskipun sentimen negatif mendominasi, besarnya persentase opini positif menunjukkan bahwa masih terdapat potensi untuk membangun reputasi yang lebih baik. Upaya komunikasi yang lebih efektif, peningkatan layanan pemerintah yang dapat meningkatkan kepercayaan, serta respons yang lebih proaktif terhadap kritik dapat menjadi strategi yang tepat untuk meningkatkan persepsi publik terhadap Danantara. Dengan memahami pola sentimen yang berkembang, pemerintah dapat mengambil langkah yang lebih terarah dalam membangun interaksi yang lebih positif dengan masyarakat.

4.2. Exploratory Data Analysis

Bagian ini menunjukkan detail dari nilai polaritas sentimen sampai dengan hubungan sentimen masyarakat dengan indikator dalam media sosial *Twitter* seperti *retweet* maupun *reply*.

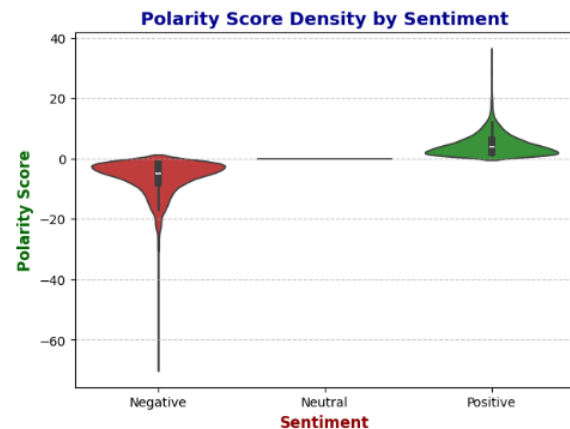


Gambar 5. Distribusi Skor Polaritas Sentimen

Distribusi skor polaritas sentimen terhadap Danantara pada Gambar 5 menunjukkan pola berbentuk distribusi normal dengan puncak yang berada di sekitar nol. Walaupun begitu dapat terlihat bahwa terdapat variasi ke arah sentimen positif dan negatif. Rentang polaritas yang menyebar ke kiri dan kanan menunjukkan adanya perbedaan opini yang cukup besar, di mana beberapa pengguna memiliki sentimen yang sangat negatif, sementara yang lain menunjukkan dukungan yang kuat terhadap Danantara.

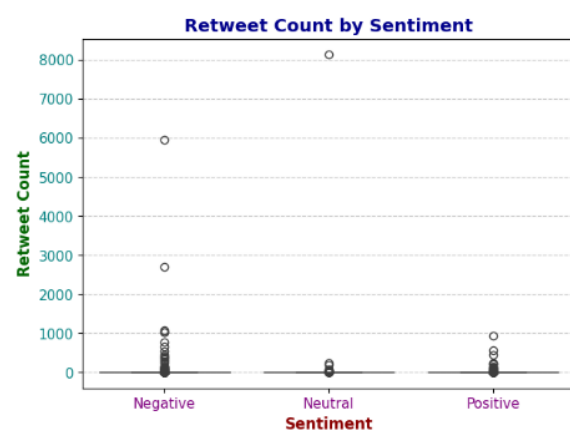
Meskipun distribusi terlihat simetris, terdapat sedikit kecenderungan ke arah negatif, sebagaimana ditunjukkan oleh lebih banyak data yang tersebar di

sisi kiri sumbu nol. Hal ini sejalan dengan distribusi sentimen sebelumnya, di mana opini negatif lebih dominan dibandingkan positif. Dengan demikian, analisis polaritas ini mengonfirmasi bahwa terdapat perbedaan signifikan dalam persepsi masyarakat terhadap Danantara, dengan sebagian besar perbincangan mencerminkan ketidakpuasan atau kekhawatiran terhadap kebijakan pemerintah Indonesia terkait dibentuknya Danantara ini.



Gambar 6. Violin Plot Distribusi Polaritas Sentimen

Visualisasi violin plot pada Gambar 6 semakin menegaskan distribusi polaritas sentimen terhadap Danantara, di mana sentimen negatif memiliki sebaran yang lebih luas dengan beberapa titik ekstrem yang menunjukkan skor polaritas yang sangat rendah. Hal ini menandakan adanya kelompok yang memiliki opini sangat negatif terhadap Danantara. Sebaliknya, sentimen positif memiliki distribusi yang lebih terpusat di sekitar skor polaritas yang lebih tinggi, tetapi dengan variasi yang lebih kecil dibandingkan sentimen negatif. Dengan demikian, analisis ini mengonfirmasi bahwa meskipun ada sentimen positif terhadap Danantara, dominasi opini negatif tetap lebih signifikan dan lebih bervariasi dalam intensitasnya.

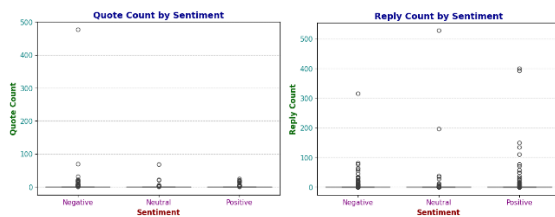


Gambar 7. Persebaran Jumlah Retweet

Berdasarkan visualisasi jumlah *retweet* berdasarkan sentimen seperti pada Gambar 7, dapat dilihat bahwa *tweet* dengan sentimen negatif

cenderung memiliki jumlah *retweet* yang lebih tinggi dibandingkan dengan sentimen netral maupun positif. Hal ini terlihat dari adanya beberapa *outlier* pada kategori sentimen negatif yang mencapai lebih dari 6000 hingga 8000 *retweet*. Banyaknya *retweet* pada *tweet* negatif ini menunjukkan bahwa opini negatif mengenai Danantara lebih sering disebarluaskan, kemungkinan besar karena adanya ketertarikan publik yang lebih tinggi terhadap isu-isu yang mengandung kritik atau kontroversi.

Sementara itu, sentimen netral dan positif memiliki jumlah *retweet* yang jauh lebih rendah dan cenderung lebih merata tanpa *outlier* yang mencolok. Meskipun terdapat beberapa *tweet* dengan sentimen positif yang mendapatkan lebih banyak *retweet*, distribusinya masih lebih kecil dibandingkan dengan sentimen negatif. Hal ini mengindikasikan bahwa opini positif tentang Danantara mungkin tidak sebesar opini negatif dalam hal penyebaran di media sosial. Fenomena ini sejalan dengan tren umum di media sosial, di mana konten yang memicu emosi, terutama kemarahan atau kekecewaan, lebih mudah viral dibandingkan dengan konten yang bersifat netral atau mendukung.

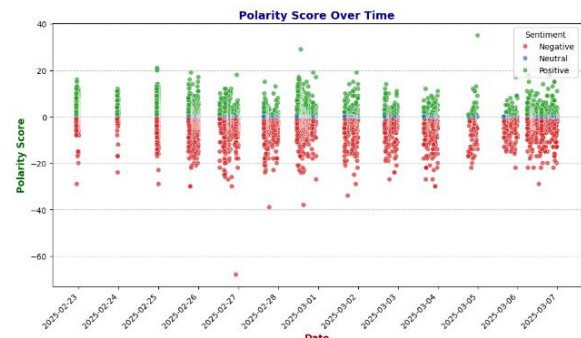


Gambar 8. Persebaran Jumlah Quote dan Reply

Berdasarkan analisis jumlah *quote* dan *reply* terhadap sentimen tentang Danantara pada Gambar 8, terlihat bahwa sentimen negatif memiliki jumlah *quote* tertinggi dibandingkan dengan sentimen netral dan positif. Ini menunjukkan bahwa banyak pengguna yang mengutip *tweet* dengan sentimen negatif untuk memberikan opini tambahan, baik dalam bentuk persetujuan maupun bantahan. Beberapa *outlier* pada kategori negatif menunjukkan jumlah *quote* yang sangat tinggi, yang menandakan bahwa diskusi mengenai kritik terhadap Danantara cukup aktif di Twitter. Sentimen netral dan positif juga memiliki beberapa *tweet* yang dikutip dalam jumlah sedang, tetapi tidak seviral *tweet* negatif.

Sementara itu, jumlah *reply* untuk masing-masing sentimen menunjukkan pola yang serupa. Sentimen negatif memiliki jumlah *reply* yang cukup tinggi, yang menunjukkan bahwa pengguna lebih banyak berdiskusi atau berdebat terkait *tweet* negatif mengenai Danantara. Beberapa *tweet* negatif memiliki *reply* di atas 500, yang menunjukkan bahwa *tweet* tersebut menimbulkan perdebatan yang cukup panjang. Sentimen positif juga memiliki beberapa *outlier* dengan jumlah *reply* yang cukup banyak, meskipun secara keseluruhan jumlahnya masih lebih rendah dibandingkan dengan sentimen negatif.

Sentimen netral memiliki jumlah *reply* yang relatif lebih sedikit, menunjukkan bahwa perbincangan di media sosial lebih didominasi oleh opini yang kuat, baik positif maupun negatif, dibandingkan dengan *tweet* yang bersifat netral.



Gambar 9. Nilai Polaritas Sentimen Setiap Tanggal

Grafik pada Gambar 9 menunjukkan distribusi skor polaritas sentimen terhadap Danantara dari tanggal 23 Februari hingga 7 Maret 2025. Terlihat bahwa pada tanggal 24 Februari, yang merupakan hari peresmian Danantara, terdapat campuran sentimen positif dan negatif dengan sebaran yang cukup luas. Meskipun terdapat *tweet* dengan polaritas tinggi yang menunjukkan dukungan terhadap Danantara, sentimen negatif juga cukup dominan dengan skor polaritas yang rendah. Ini mengindikasikan bahwa peluncuran Danantara mendapatkan perhatian yang besar di media sosial, baik dalam bentuk dukungan maupun kritik. Setelah tanggal 24, sentimen tetap bervariasi, namun tampaknya sentimen negatif semakin meningkat pada hari-hari berikutnya, yang dapat mengindikasikan adanya reaksi atau diskusi lebih lanjut mengenai peresmian Danantara.

Secara keseluruhan, pola sentimen yang terlihat menunjukkan adanya perdebatan di media sosial terkait Danantara, dengan banyak *tweet* yang berada dalam kategori negatif dan memiliki skor polaritas yang cukup rendah. Beberapa lonjakan sentimen positif terjadi dalam beberapa hari setelah peresmian, yang mungkin berasal dari diskusi adanya kepercayaan masyarakat kepada pemerintah. Namun, kehadiran sentimen negatif yang terus berlanjut mengindikasikan adanya ketidakpuasan atau kritik yang masih aktif dalam diskusi publik mengenai kebijakan pemerintah Indonesia terkait Danantara.

5. KESIMPULAN DAN SARAN

Penelitian ini menganalisis sentimen masyarakat Indonesia terhadap Danantara menggunakan pendekatan *Lexicon-Based Sentiment Analysis* pada data dari media sosial Twitter. Hasil penelitian menunjukkan bahwa sentimen publik terhadap Danantara cenderung beragam, dengan dominasi sentimen negatif (44,9%) yang sedikit lebih tinggi dibandingkan dengan sentimen positif (44,1%). Faktor utama yang mempengaruhi persepsi publik meliputi transparansi kebijakan, kredibilitas pengelolaan

investasi, serta komunikasi pemerintah terkait tujuan dan manfaat Danantara. Selain itu, sentimen netral sebesar 11% menunjukkan adanya kelompok masyarakat yang masih belum memiliki pendapat tegas terhadap keberadaan Danantara. Dibandingkan dengan model analisis sentimen berbasis *machine learning*, pendekatan *lexicon-based* menawarkan interpretasi yang lebih transparan dan efisien dalam klasifikasi sentimen, meskipun tetap memiliki keterbatasan dalam menangkap konteks kalimat yang kompleks. Secara keseluruhan, penelitian ini memberikan gambaran tentang bagaimana opini publik dapat digunakan sebagai indikator awal untuk mengukur penerimaan masyarakat terhadap kebijakan ekonomi strategis yang diterapkan oleh pemerintah.

Berdasarkan hasil penelitian ini, terdapat beberapa rekomendasi untuk penelitian selanjutnya. Pertama, perlu dilakukan eksplorasi terhadap model *hybrid* yang menggabungkan pendekatan *lexicon-based* dengan *machine learning* atau *large language models* (LLM) untuk meningkatkan akurasi klasifikasi sentimen, terutama dalam menangkap konteks, sarkasme, dan ironi dalam teks. Kedua, penelitian mendatang dapat memperluas cakupan data dengan mempertimbangkan berbagai platform media sosial lainnya, seperti Facebook dan Instagram, guna memperoleh perspektif yang lebih komprehensif. Ketiga, analisis faktor eksternal, seperti pemberitaan media dan respons pejabat terkait, dapat dikombinasikan untuk memahami lebih dalam dinamika sentimen publik. Terakhir, pengembangan kamus sentimen yang lebih spesifik terhadap domain ekonomi dan investasi di Indonesia dapat meningkatkan keakuratan analisis sentimen dalam studi-studi mendatang.

DAFTAR PUSTAKA

- [1] L. Lie, "Danantara: An SOE Superholding à la Temasek?," 2024.
- [2] Pemerintah Republik Indonesia, "Peraturan Pemerintah Republik Indonesia Nomor 10 Tahun 2025 Tentang Organisasi dan Tata Kelola Badan Pengelola Investasi Daya Anagata Nusantara," 2025.
- [3] H. Nguyen and S. Gokhale, "An efficient approach to identifying antigovernment sentiment on Twitter during Michigan protests," *PeerJ Comput Sci*, vol. 8, 2022, doi: 10.7717/PEERJ-CS.1127.
- [4] A. Borg and M. Boldt, "Using VADER sentiment and SVM for predicting customer response sentiment," *Expert Syst Appl*, vol. 162, p. 113746, 2020, doi: 10.1016/j.eswa.2020.113746.
- [5] M. Isnan, G. N. Elwirehardja, and B. Pardamean, "Sentiment Analysis for TikTok Review Using VADER Sentiment and SVM Model," *Procedia Comput Sci*, vol. 227, pp. 168–175, 2023, doi: 10.1016/j.procs.2023.10.514.
- [6] M. H. Hoti and J. Ajdari, "Sentiment Analysis Using the Vader Model for Assessing Company Services Based on Posts on Social Media," *SEEU Review*, vol. 18, no. 2, pp. 19–33, 2023, doi: 10.2478/seeur-2023-0043.
- [7] A. M. Van Der Veen and E. Bleich, "The advantages of lexicon-based sentiment analysis in an age of machine learning," *PLoS One*, vol. 20, no. 1 January, pp. 1–19, 2025, doi: 10.1371/journal.pone.0313092.
- [8] M. E. Mowlaei, M. Saniee Abadeh, and H. Keshavarz, "Aspect-based sentiment analysis using adaptive aspect-based lexicons," *Expert Syst Appl*, vol. 148, p. 113234, 2020, doi: 10.1016/j.eswa.2020.113234.
- [9] W. Megginson and V. Fotak, "Government equity investments in coronavirus bailouts: Why, how, when?," *Journal of Law, Finance, and Accounting*, vol. 6, no. 1, pp. 1–49, 2021, doi: 10.1561/108.00000050.
- [10] A. James, T. Retting, J. F. Shogren, B. Watson, and S. Wills, "Sovereign Wealth Funds in Theory and Practice," *Annu Rev Resour Economics*, vol. 14, pp. 621–646, 2022, doi: 10.1146/annurev-resource-111920-015758.
- [11] L. Nawo and H. Njangang, "The effect of covid-19 outbreak on foreign direct investment: do sovereign wealth funds matter?," *Transnational Corporations Review*, vol. 14, no. 1, pp. 1–17, 2022, doi: 10.1080/19186444.2021.1964313.
- [12] K. E. Naresh Kumar and V. Uma, "Intelligent sentiment-based lexicon for context-aware sentiment analysis: optimized neural network for sentiment classification on social media," *Journal of Supercomputing*, vol. 77, no. 11, pp. 12801–12825, 2021, doi: 10.1007/s11227-021-03709-4.
- [13] R. Obiedat *et al.*, "Sentiment Analysis of Customers' Reviews Using a Hybrid Evolutionary SVM-Based Approach in an Imbalanced Data Distribution," *IEEE Access*, vol. 10, pp. 22260–22273, 2022, doi: 10.1109/ACCESS.2022.3149482.
- [14] A. Bayhaqy, "Analisa Sentimen Dengan Pelabelan Otomatis Menggunakan VADER Tentang Islamophobia Pada Media Sosial," no. August 2019, pp. 9–11, 2020, doi: 10.13140/RG.2.2.19304.08966.
- [15] Y. Nooryuda Prasetya and D. Winarso, "Penerapan Lexicon Based Untuk Analisis Sentimen Pada Twitter Terhadap Isu Covid-19," *Jurnal Fasilkom*, vol. 11, no. 2, pp. 97–103, 2021, [Online]. Available: <https://ejurnal.umri.ac.id/index.php/JIK/article/view/2772/1566>
- [16] A. Erkan and T. Gungor, "Analysis of Deep Learning Model Combinations and Tokenization Approaches in Sentiment Classification," *IEEE Access*, vol. 11, no. November, pp. 134951–134968, 2023, doi: 10.1109/ACCESS.2023.3337354.

- [17] A. Jabbar, S. Iqbal, M. I. Tamimy, A. Rehman, S. A. Bahaj, and T. Saba, "An Analytical Analysis of Text Stemming Methodologies in Information Retrieval and Natural Language Processing Systems," *IEEE Access*, vol. 11, no. October, pp. 133681–133702, 2023, doi: 10.1109/ACCESS.2023.3332710.
- [18] S. Sarica and J. Luo, "Stopwords in technical language processing," *PLoS One*, vol. 16, no. 8 August, pp. 1–13, 2021, doi: 10.1371/journal.pone.0254937.
- [19] Fajri Koto and Gemala Y. Rahmaningtyas, "InSet Lexicon: Evaluation of a Word List for Indonesian Sentiment Analysis in Microblogs," *2017 International Conference on Asian Language Processing (IALP)*, pp. 391–394, 20AD