

---

# Academia Open



*By Universitas Muhammadiyah Sidoarjo*

---

# Academia Open

Vol. 11 No. 1 (2026): June  
DOI: 10.21070/acopen.11.2026.14210

## Table Of Contents

<b>Journal Cover</b> .....	1
<b>Author[s] Statement</b> .....	3
<b>Editorial Team</b> .....	4
<b>Article information</b> .....	5
Check this article update (crossmark) .....	5
Check this article impact .....	5
Cite this article.....	5
<b>Title page</b> .....	6
Article Title .....	6
Author information .....	6
Abstract .....	6
<b>Article content</b> .....	8

## Originality Statement

The author[s] declare that this article is their own work and to the best of their knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the published of any other published materials, except where due acknowledgement is made in the article. Any contribution made to the research by others, with whom author[s] have work, is explicitly acknowledged in the article.

## Conflict of Interest Statement

The author[s] declare that this article was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Copyright Statement

Copyright © Author(s). This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

# Academia Open

Vol. 11 No. 1 (2026): June  
DOI: 10.21070/acopen.11.2026.14210

## EDITORIAL TEAM

### Editor in Chief

Mochammad Tanzil Multazam, Universitas Muhammadiyah Sidoarjo, Indonesia

### Managing Editor

Bobur Sobirov, Samarkand Institute of Economics and Service, Uzbekistan

### Editors

Fika Megawati, Universitas Muhammadiyah Sidoarjo, Indonesia

Mahardika Darmawan Kusuma Wardana, Universitas Muhammadiyah Sidoarjo, Indonesia

Wiwit Wahyu Wijayanti, Universitas Muhammadiyah Sidoarjo, Indonesia

Farkhod Abdurakhmonov, Silk Road International Tourism University, Uzbekistan

Dr. Hindarto, Universitas Muhammadiyah Sidoarjo, Indonesia

Evi Rinata, Universitas Muhammadiyah Sidoarjo, Indonesia

M Faisal Amir, Universitas Muhammadiyah Sidoarjo, Indonesia

Dr. Hana Catur Wahyuni, Universitas Muhammadiyah Sidoarjo, Indonesia

Complete list of editorial team ([link](#))

Complete list of indexing services for this journal ([link](#))

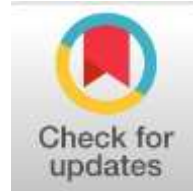
How to submit to this journal ([link](#))

# Academia Open

Vol. 11 No. 1 (2026): June  
DOI: 10.21070/acopen.11.2026.14210

## Article information

**Check this article update (crossmark)**



**Check this article impact (\*)**



**Save this article to Mendeley**



(\*) Time for indexing process is various, depends on indexing database platform

## Danantara YouTube Sentiment Shows Public Transparency Concerns: Sentimen YouTube Danantara Menunjukkan Kekhawatiran Transparansi Publik

**Waraney Vincent Beckham Wahani, vincentwaraney@gmail.com (\*)**

*Program Studi Teknik Informatika, Universitas Negeri Manado, Indonesia*

**Alfiansyah Hasibuan, alfiansyahhasibuan@unima.ac.id**

*Program Studi Teknik Informatika, Universitas Negeri Manado, Indonesia*

**Medi Hermanto Tinambunan, meditinambunan@unima.ac.id**

*Program Studi Teknik Informatika, Universitas Negeri Manado, Indonesia*

(\*) Corresponding author

### Abstract

**General Background** Social media comments offer valuable data for analyzing public discourse on policy issues. **Specific Background** This study investigates YouTube comments about Danantara, Indonesia's strategic investment body, using Natural Language Processing with 7,294 comments. **Knowledge Gap** Previous studies often analyze sentiment and topics separately, without integrated analysis or iterative labeling. **Aims** The study aims to classify sentiment using Support Vector Machine (SVM) and identify topics with Latent Dirichlet Allocation (LDA). **Results** 74.9% accuracy was achieved with SVM, classifying 58.0% of comments as negative, 29.3% neutral, and 12.8% positive. LDA revealed 6 topics for neutral, 4 for positive, and 3 for negative sentiment, with key concerns about transparency and corruption. **Novelty** This study integrates SVM and LDA with Human in the Loop labeling to capture both sentiment and topic substance. **Implications** Findings offer insights for improving transparency and public communication, while contributing to text mining in digital discourse.

### Highlights

- The classifier achieved 74.9% accuracy after Human in the Loop labeling and manual verification.
- Unfavorable polarity reached 58.0%, followed by neutral at 29.3% and positive at 12.8%.
- Coherence scores selected 6 neutral, 4 positive, and 3 critical thematic clusters.

### Keywords

Danantara; Sentiment Analysis; Topic Modeling; Support Vector Machine; Latent Dirichlet Allocation

Published date: 2026-05-12

## PENDAHULUAN

Perkembangan teknologi informasi telah mendorong media sosial menjadi ruang diskusi publik digital. YouTube tidak hanya berfungsi sebagai platform berbagi video, tetapi juga menyediakan kolom komentar yang memungkinkan pengguna menyampaikan opini secara langsung terhadap suatu isu. Komentar-komentar tersebut merepresentasikan persepsi publik dan menghasilkan data teks dalam jumlah besar yang bersifat tidak terstruktur [1].

Danantara adalah lembaga investasi strategis pemerintah untuk mengoptimalkan aset negara dan mendorong pertumbuhan ekonomi [2]. Kehadirannya memicu beragam respons publik di YouTube terkait transparansi dan dampak ekonomi.

Karakter komentra yang besar dan informal membuat analisis manual tidak efisien, sehingga diperlukan pendekatan text mining dan Natural Language Processing (NLP) untuk analisis sistematis [3].

Analisis sentimen telah banyak digunakan untuk mengidentifikasi kecenderungan opini publik pada media sosial terhadap suatu isu atau kebijakan. Di sisi lain, pemodelan topik dimanfaatkan untuk mengungkap tema-tema utama yang berkembang dalam kumpulan teks tidak terstruktur. Kedua pendekatan tersebut memiliki fokus analisis yang berbeda namun saling melengkapi. Analisis sentimen memberikan gambaran mengenai arah opini, sedangkan pemodelan topik membantu mengidentifikasi substansi atau isu yang dibahas [2],[3]. Oleh karena itu, penggabungan kedua pendekatan dalam satu kerangka analisis dapat memberikan pemahaman yang lebih komprehensif terhadap pola diskusi publik dalam ruang digital.

Analisis sentimen dalam penelitian ini dilakukan menggunakan Support Vector Machine (SVM), yaitu algoritma supervised learning yang bekerja dengan mencari hyperplane optimal untuk memisahkan data ke dalam kelas tertentu berdasarkan margin maksimum [4]. SVM dikenal efektif dalam klasifikasi teks karena mampu menangani data berdimensi tinggi dan bersifat sparse, seperti representasi fitur berbasis TF-IDF [4]. Untuk melengkapi analisis polaritas opini, penelitian ini juga menerapkan pemodelan topik menggunakan Latent Dirichlet Allocation (LDA) [5], sebuah metode unsupervised learning berbasis probabilistik yang memodelkan setiap dokumen sebagai distribusi atas sejumlah topik laten dan setiap topik sebagai distribusi probabilitas kata [6].

Meskipun berbagai penelitian sebelumnya telah mengkaji analisis sentimen maupun pemodelan topik secara terpisah, masih terdapat keterbatasan dalam mengintegrasikan kedua pendekatan tersebut secara simultan pada konteks isu kebijakan publik di Indonesia, khususnya yang bersumber dari data komentar YouTube. Selain itu, sebagian penelitian masih bergantung pada pelabelan data secara otomatis atau sepenuhnya manual tanpa mekanisme validasi iteratif, sehingga berpotensi menurunkan kualitas hasil klasifikasi.

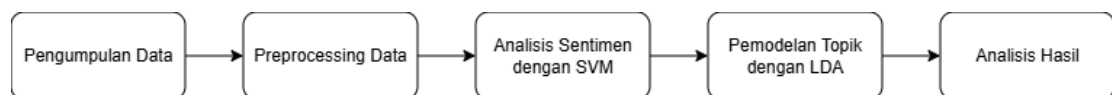
Penelitian ini menghadirkan kebaruan melalui integrasi SVM dan LDA untuk analisis sentiment dan topik secara terpadu, serta penggunaan Human in the Loop guna meningkatkan akurasi pelabelan. Fokus pada Danantara juga menjadi kontribusi kontekstual, mengingat kajian persepsi publik terkait kebijakan investasi pemerintah di media social masih terbatas.

## METODE

### A. Tahap Penelitian

Penelitian ini meliputi pengumpulan dan pelabelan komentar YouTube terkait Sanantara, pra-proses teks, pembobotan TF-IDF, analisis menggunakan SVM dan LDA, serta tahap evaluasi. Alur penelitian disajikan pada Gambar 1.

Gambar 1 . Tahapan Penelitian



### B. Pengumpulan Data

Penelitian ini menggunakan data sekunder berupa komentar pengguna YouTube pada tayangan terkait Danantara. Data dikumpulkan melalui proses scraping menggunakan YouTube Data API dari delapan video yang telah ditentukan berdasarkan relevansi konten terhadap isu Danantara. Data yang akan dikumpulkan berupa teks komentar.

### C. Preprocessing Teks

Preprocessing merupakan tahap awal yang mencakup serangkaian teknik untuk membersihkan dan menyiapkan data mentah sebelum proses analisis atau pemodelan. Tahap ini bertujuan meningkatkan kualitas data dengan mengurangi artefak atau noise tertentu sehingga hasil prediksi menjadi lebih akurat [7]. Berikut langkah-langkah pada tahap ini :

#### 1. Case Folding

Mengubah seluruh huruf dalam teks menjadi huruf kecil untuk menyeragamkan format penulisan dan menghindari duplikasi kata akibat perbedaan kapitalisasi.

## 2. Data Cleaning

Menghapus karakter non-alfabet, tanda baca, angka, URL, serta simbol yang tidak relevan agar teks lebih bersih dan siap diproses.

## 3. Tokenization

Memecah teks menjadi unit-unit kata (token) berdasarkan spasi atau delimiter tertentu sehingga setiap kata dapat dianalisis secara individual.

## 4. Normalisasi Kata

Mengubah kata tidak baku, singkatan, dan bahasa slang yang umum digunakan dalam komentar YouTube menjadi bentuk baku untuk mengurangi variasi kosakata dengan makna serupa.

## 5. Stopword Removal

Menghapus kata-kata umum yang memiliki nilai informasi rendah, seperti kata hubung dan kata depan, agar analisis lebih fokus pada kata yang informatif.

## D. Support Vector Machine

Support Vector Machine (SVM) adalah metode supervised learning yang efektif untuk klasifikasi data teks kompleks [10]. Penelitian ini menggunakan kernel linear dengan parameter C bernilai 1. SVM bekerja dengan menemukan hyperlane optimal sebagai pemisah terbaik antar kelas data [11], sehingga mampu membedakan dua kelompok secara efektif [12],[13].

## E. Latent Dirichlet Allocation

Latent Dirichlet Allocation (LDA) merupakan Teknik pemodelan topik yang digunakan untuk menemukan pola dalam Kumpulan dokumen, sehingga dokumen dapat dikelompokkan secara otomatis ke dalam beberapa topik berbeda [14],[15].

# HASIL DAN PEMBAHASAN

## A. Pengumpulan Data

Pengumpulan data dilakukan melalui proses scraping menggunakan YouTube Data API yang dijalankan pada platform Google Colab. Scraping dilakukan terhadap komentar pengguna pada delapan video YouTube yang membahas Danantara. Data yang diperoleh kemudian disimpan dalam format CSV untuk proses analisis lebih lanjut. Total data yang berhasil dikumpulkan sebanyak 7.294 komentar. Dalam penelitian ini, atribut yang digunakan adalah teks komentar sebagai objek utama analisis.

Pelabelan dilakukan dengan pendekatan Human in the Loop pada 5.000 komentar menggunakan Gemini 2.5 Flash dan diverifikasi manual. Hasilnya didominasi sentiment negatif (3.260), diikuti netral (1.047) dan positif (693), yang mencerminkan kondisi opini publik saat pengambilan sampel acak. Selanjutnya, seluruh dataset tersebut dialokasikan ke dalam data latih dan data uji dengan rasio pembagian 80:20, di mana 80% (4.000 data) digunakan untuk melatih algoritma dan 20% (1.000 data) sisanya digunakan untuk mengevaluasi performa model klasifikasi sentimen yang dihasilkan. Tabel 1 menampilkan sampel data mentah yang telah melalui proses pelabelan manual.

**Tabel 1 . Sampel data mentah dan label**

Text	Label
ini bagus sih, sudah saatnya kita menjadi Investor. kalau bisa ajak orang kaya-orang kaya di negeri ini buat berani menjadi investor luar negeri. seperti Djarum misalnya, yang sudah mulai berani investasi ke luar negeri. respect	Positif
Tidak ada dalam sejarah Indonesia lembaga pengelola keuangan ber-ending baik. Jiwasraya, taspen, asabri, dll. Apalagi digagas oleh seorang "mantan" koruptor. Waktunya dibawa senang, sebelum Indonesia bubar 🙄	Negatif
Bagaimana hub danantata dengan kementeeian keuangan	Netral

## B. Preprocessing Teks

Untuk menghasilkan data yang bersih dan konsisten sebelum masuk ke tahap klasifikasi SVM dan pemodelan topik LDA, dilakukan lima tahap preprocessing dan hasilnya dapat dilihat pada Tabel 2.

**Tabel 2. Hasil Preprocessing Data**

Komponen	Text
Text Asli	Wah bagaimana ini??? Idenya bagus tapi aplikasinya bagaimana nih?? Kalo semua deviden dikumpulkan jadi satu tdk masuk RAPBN dimasukkan DANANTARA, kalo dikorup semua bagaimana? Kalo diaplikasikan sih ga pp 🤔🤔🤔
Case Folding	wah bagaimana ini??? idenya bagus tapi aplikasinya bagaimana nih?? kalo semua deviden dikumpulkan jadi satu tdk masuk rapbn dimasukkan danantara, kalo dikorup semua

	bagaimana? kalo diaplikasikan sih ga pp 🤔🤔🤔
<b>Data Cleaning</b>	wah bagaimana ini idenya bagus tapi aplikasinya bagaimana nih kalo semua deviden dikumpulkan jadi satu tdk masuk rapbn dimasukkan danantara kalo dikorup semua bagaimana kalo diaplikasikan sih ga pp
<b>Tokenization</b>	['wah', 'bagaimana', 'ini', 'idenya', 'bagus', 'tapi', 'aplikasinya', 'bagaimana', 'nih', 'kalo', 'semua', 'deviden', 'dikumpulkan', 'jadi', 'satu', 'tdk', 'masuk', 'rapbn', 'dimasukkan', 'danantara', 'kalo', 'dikorup', 'semua', 'bagaimana', 'kalo', 'diaplikasikan', 'sih', 'ga', 'pp']
<b>Normalisasi</b>	['wah', 'bagaimana', 'ini', 'idenya', 'bagus', 'tapi', 'apliaksinya', 'bagaimana', 'ini', 'kalau', 'semua', 'deviden', 'dikumpulkan', 'jadi', 'satu', 'tidak', 'masuk', 'rapbn', 'dimasukkan', 'danantara', 'kalau', 'dikorup', 'semua', 'bagaimana', 'kalau', 'diaplikasikan', 'sih', 'tidak', 'apa-apa']
<b>Stopword Removal</b>	['wah', 'bagaimana', 'idenya', 'bagus', 'aplikasinya', 'bagaimana', 'kalau', 'semua', 'deviden', 'dikumpulkan', 'jadi', 'satu', 'tidak', 'masuk', 'rapbn', 'dimasukkan', 'danantara', 'kalau', 'dikorup', 'semua', 'bagaimana', 'kalau', 'diaplikasikan', 'tidak', 'apa-apa']

Setelah prapemrosesan, teks yang telah bersih diubah ke bentuk numerik agar dapat diproses algoritma, menggunakan metode Term Frequency Inverse Document Frequency (TF-IDF). Metode ini tidak hanya menghitung frekuensi kemunculan sebuah kata dalam satu komentar (Term Frequency), tetapi juga mempertimbangkan seberapa penting atau unik kata tersebut di dalam seluruh dataset (Inverse Document Frequency) [8]. Kata-kata yang terlalu umum akan diberikan bobot rendah, sementara kata kunci yang kuat seperti "korupsi", "untung", atau "aman" akan diberikan bobot yang lebih tinggi untuk membantu proses klasifikasi.

### C. Klasifikasi dengan Support Vector Machine

Data hasil ekstraksi fitur kemudian dimasukkan ke model Support Vector Machine (SVM) dengan kernel linear (atau RBF). SVM mengklasifikasikan data dengan menentukan hyperplane optimal untuk memisahkan tiga sentiment: positif, negatif, dan netral. Evaluasi dilakukan menggunakan confusion matrix untuk membandingkan data actual dan prediksi, menghasilkan nilai TN, FP, FN, dan TP, yang digunakan menghitung akurasi, presisi, recall, dan F1 score [17]. Implementasi menggunakan sklearn dan metrics, dengan hasil ditampilkan pada gambar 2.

**Gambar 2. Confusion Matrix**



Berdasarkan hasil pengujian menggunakan Confusion Matrix, model klasifikasi Support Vector Machine (SVM) yang dikembangkan berhasil mencapai nilai akurasi sebesar 74,9%. Angka ini menunjukkan bahwa model secara fungsional telah memiliki kemampuan yang reliabel dalam memetakan opini publik mengenai isu Danantara.

Performa terbaik terlihat pada sentiment negatif, dengan 547 dari 652 sata berhasil diprediksi benar (Precision 0,85; Recall 0,84), menunjukkan kemampuan tinggi dalam menangkap pola kritik dan kekhawatiran. Kesalahan klasifikasi tergolong rendah, dengan 82 data terbaca netral dan 23 positif.

Pada sentiment positif, model cukup akurat dengan precision 0,66 dan recall 0,63, meski masih dipengaruhi kemiripan diksi. Sementara itu, pada kelas netral, 114 dari 209 data terklasifikasi benar, sedangkan 72 data keliru sebagai akibat ambiguitas Bahasa pada komentar informatif.

Hal ini merupakan implikasi dari fenomena distribusi data di mana sentimen Negatif (support: 652) jauh lebih mendominasi dataset dibandingkan kelas Netral (209) dan Positif (139). Secara sosiologis, hal ini merefleksikan dinamika audiens YouTube yang cenderung lebih vokal dalam menyampaikan kritik dan kewaspadaan terhadap kebijakan pengelolaan investasi besar, sehingga model mendapatkan lebih banyak referensi pembelajaran pada pola kalimat negatif dibandingkan pola kalimat apresiatif atau netral.

Temuan ini sejalan dengan studi sebelumnya bahwa SVM stabil dan efektif untuk klasifikasi teks media social. Berbeda dari penelitian terdahulu, studi ini menggabungkan analisis sentiment dan topik sehingga mengungkap opini sekaligus konteksnya.

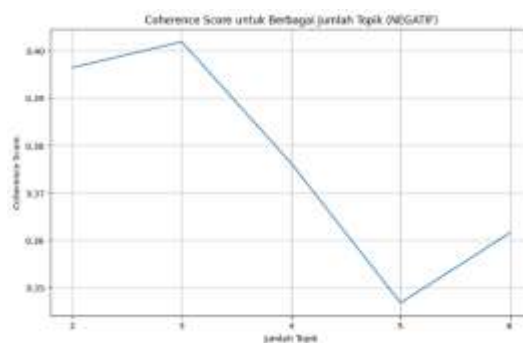
Setelah melalui tahap evaluasi performa menggunakan Confusion Matrix dan memastikan bahwa model SVM memiliki tingkat reliabilitas yang memadai pada data uji, langkah selanjutnya adalah melakukan implementasi model terhadap sisa





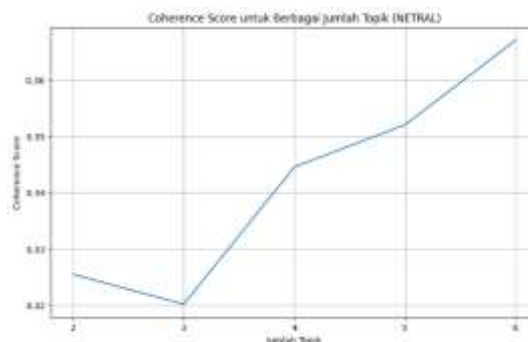
Gambar 8 menunjukkan grafik skor koherensi untuk sentimen positif, di mana nilai optimal berada pada 4 topik dengan skor tertinggi mencapai 0,4341. Setelah mencapai titik puncak tersebut, skor koherensi mengalami penurunan pada 5 topik dan berlanjut hingga 6 topik.

**Gambar 9.** Grafik Skor Koherensi Sentimen Negatif



Gambar 9 Grafik Skor Koherensi Sentimen Negatif menunjukkan bahwa nilai optimal berada pada 3 topik dengan skor tertinggi mencapai 0,4019. Setelah mencapai titik tersebut, grafik menunjukkan tren penurunan yang signifikan hingga mencapai nilai terendah pada 5 topik, sebelum akhirnya mengalami sedikit kenaikan pada 6 topik.

**Gambar 10.** Grafik Skor Koherensi Sentimen Netral



Gambar 10 Grafik Skor Koherensi Sentimen Netral menunjukkan tren yang cenderung naik seiring bertambahnya jumlah topik, dengan nilai tertinggi mencapai 0,3671 pada 6 topik. Pola ini mengindikasikan bahwa jumlah topik tersebut merupakan yang paling ideal untuk kategori ini dibandingkan dengan jumlah topik sebelumnya.

Berdasarkan hasil analisis grafik skor koherensi, penentuan jumlah topik optimal dilakukan dengan mengidentifikasi skor tertinggi di setiap kategori sentimen guna menjamin keterkaitan makna yang paling kuat. Sentiment negatif paling konsisten dengan 3 topik (0,4019), positif optimal pada 4 topik (0,4341), dan netral pada 6 topik (0,3671), mencerminkan perbedaan kompleksitas respons publik.

Hasil ini menegaskan efektivitas LDA dalam mengungkap topik laten serta menunjukkan bahwa jumlah topik optimal berbeda tiap subset. Pemisahan berdasarkan sentiment juga memperkaya analisis dibanding pendekatan Tunggal. Setelah menetapkan jumlah topik optimal berdasarkan skor koherensi tertinggi di setiap kategori, tabel berikut menyajikan interpretasi mendalam terhadap kumpulan kata kunci yang merepresentasikan substansi opini publik pada masing-masing sentimen :

**Tabel 4.** Interpretasi Topik setiap sentiment

Sentimen	Topik	Kata Kunci Utama	Interpretasi Sementara
<b>POSITIF</b>	0	bagus, sangat, mudah, terimakasih, danantara, ferry, banget, penjelasan, semoha, sekali	Apresiasi terhadap kejelasan penjelasan dan kualitas konten.
	1	kalau, lebih, orang, buat, selalu, terus, suka, video, jelas, bukan	Kepuasan audiens terhadap kejelasan narasi dalam video.
	2	baik, indonesia, kalua, banget, bikin, perusahaan, segmen, swasta, kerja, akhirnya	Dukungan terhadap keterlibatan sektor swasta dan perusahaan di Indonesia.
	3	kasih, terima, bahasa, bayi, semoga, semangat, banyak, paham, pekerja, konten	Ungkapan terima kasih dan harapan positif bagi para pekerja serta konten edukasi.
<b>NETRAL</b>	0	kerja, baik, main, benar, coba, sama, kalua, belum, banyak, jelas.	Diskusi mengenai mekanisme kerja dan pembuktian system.
	1	kalau, pos, berarti, pegawai, kadang, padahal, bg, pt mitra, sama	Pembahasan mengenai status pegawai dan kemitraan lembaga (PT/Pos)
	2	danantara, bumh, investasi, mau, sama, kan, baik, kalua, asset prabowo	Analisis teknis mengenai pengelolaan asset Danantara dan BUMN.
	3	danantara, negara, bayi, bahasa, indonesia, the, tahun, bahas, terlalu, close	Diskusi umum mengenai keberadaan Danantara dalam konteks negara.
	4	mitra, pos, semua, buka, channel, terus, orang, indonesia, kurir, kerja	Pembahasan mengenai peran kurir dan kemitraan dalam ekosistem kerja.
	5	dana, pertiwi, si, uang, investasi, kalau, danantara, sama, kayak, orang	Tanya jawab objektif seputar aliran dana investasi dan entitas terkait.
<b>NEGATIF</b>	0	bumh, sama, danantara, kalau, benar, orang, mau, mdb, indonesia, malaysia	Kekhawatiran akan perbandingan kasus hukum/finansial luar negeri (IMDB)
	1	negara, kalau, korupsi, rakyat, orang, banyak koruptor, indonesia, semua, pemerintah	Kritik tajam terhadap pemerintah terkait potensi korupsi dan Nasib rakyat.
	2	dana, pertiwi, kalau, si, danantara, korupsi, uang, sama, negara, investasi	Skeptisime terhadap keamanan pengelolaan dana dan aset negara.

## KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan, dapat disimpulkan bahwa:

1. Akurasi Klasifikasi: Model Support Vector Machine (SVM) berhasil mengklasifikasikan sentimen publik dengan tingkat akurasi sebesar 74,9%, yang menunjukkan reliabilitas model dalam memetakan persepsi masyarakat terhadap isu Danantara.
2. Dominasi sentiment: terjadi ketimpangan data, di mana sentiment negatif mendominasi sebesar 58%, mencerminkan sikap kritis publik terhadap kebijakan investasi pemerintah.
3. Substansi opini: hasil LDA menunjukkan sentiment negatif didominasi isu korupsi, perbandingan kasus hukum, dan keamanan data, sedangkan sentiment positif lebih menyoroti kejelasan informasi serta harapan bagi pekerja dan perusahaan.
4. Optimalisasi model: berdasarkan skor koherensi, sentiment netral memiliki struktur topik paling kompleks (6 topik) dibandingkan positif (4) dan negatif (3), yang menunjukkan perbedaan kedalaman pembahasan pada tiap kategori. Secara praktis, penelitian ini dapat menjadi bahan evaluasi pemerintah dalam meningkatkan transparansi, komunikasi publik, dan pengelolaan isu, serta referensi bagi media dalam Menyusun konten edukatif yang lebih efektif. Secara teoritis, penelitian ini berkontribusi pada pengembangan text mining melalui integrasi SVM dan LDA untuk mengungkap polaritas serta substansi wacana pada data teks. Penelitian selanjutnya disarankan menggunakan algoritma yang lebih kompleks seperti LSTM atau Transformer, memperluas sumber data, menerapkan Teknik penanganan data imbalance seperti oversampling atau SMOTE, serta melakukan studi longitudinal guna melihat dinamika opini publik.

## UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada seluruh pihak yang telah mendukung penelitian ini, khususnya pembimbing, institusi penyedia fasilitas dan data, serta keluarga dan rekan atas dukungan dan motivasinya. Semoga penelitian ini bermanfaat bagi pengembangan ilmu pengetahuan.

## References

- [1] F. A. Abdullah, G. N. Bakry, dan S. F. Z. Fibaeti, "Analisis Sentimen dan Jaringan Pada Video Youtube Malaka 'Memahami Danantara dengan Mudah,'" *MUKASI J. Ilmu Komun.*, vol. 4, no. 3, hal. 750–761, 2025, doi: 10.54259/mukasi.v4i3.5162.
- [2] "Danantara Indonesia," Badan Pengelola Investasi Daya Anagata Nusantara (Danantara Indonesia). Diakses: 25 Februari 2026. [Daring]. Tersedia pada: <https://www.danantaraindonesia.co.id/>
- [3] M. Fazri dan A. Voutama, "ANALISIS SENTIMEN PUBLIK TERHADAP DANANTARA DI MEDIA SOSIAL X MENGGUNAKAN NLP DAN PEMBELAJARAN MESIN," *JOISIE (Journal Inf. Syst. Informatics Eng.*, vol. 6, no. 1, hal. 197–206, 2025, doi: 10.55382/jurnalpustakaai.v5i2.1180.
- [4] D. F. Sarah, U. Khaira, dan M. F. Putri, "Analisis Sentimen Aplikasi ShopeePay Menggunakan Naïve Bayes Dan Pemodelan Topik Latent Dirichlet Allocation," *Djtechno J. Teknol. Inf.*, vol. 6, no. 2, hal. 402–416, 2025, doi: 10.46576/djtechno.v6i2.6586.
- [5] M. R. Nur, Y. Wibisono, dan R. Megasari, "Analisis Sentimen dan Pemodelan Topik pada Post tentang Merek Teknologi di X Menggunakan Fine-tuning IndoBERT dan BERTopic," *J. Komput. Teknol. Inf. Sist. Inf.*, vol. 4, no. 2, hal. 743–750, 2025, doi: 10.62712/juktisi.v4i2.508.
- [6] M. L. Hermanto, O. Ardhillah, Nky, dan A. Ibrahim, "ANALISIS SENTIMEN MASYARAKAT TERHADAP DANANTARA DI PLATFORM X DENGAN METODE SVM," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 4, hal. 6779–6785, 2025.  
[ISSN 2714-7444 \(online\)](https://acopen.umsida.ac.id), <https://acopen.umsida.ac.id>, published by [Universitas Muhammadiyah Sidoarjo](https://www.umsida.ac.id)

- [7] D. H. Rahman, Z. D. A. Alistin, dan S. Pramana, "Analisis Sentimen Dan Pemodelan Topik Pada Tweet Terkait Data Badan Pusat Statistik Tahun 2024," *Semin. Nas. Off. Stat.*, hal. 226–236, 2025, doi: 10.47080/simika.v6i2.2789.
- [8] R. Z. Firdaus, S. H. Wijoyo, dan W. Purnomo, "Analisis Sentimen Berbasis Aspek Ulasan Pengguna Aplikasi Alfagift Menggunakan Metode Random Forest dan Pemodelan Topik Latent Dirichlet Allocation," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 2, 2025, [Daring]. Tersedia pada: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/14482>
- [9] A. I. Kamil, O. N. Pratiwi, dan D. Witarsyah, "Analisis Sentimen dan Pemodelan Topik terhadap Aplikasi Pembelajaran Online pada Platform Google Play," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 10, no. 2, hal. 836–849, 2025, doi: 10.29100/jipi.v10i2.6023.
- [10] C. L. T. Boro, A. Faisol, dan D. Rudhistiar, "ANALISIS SENTIMEN TERHADAP KAMPANYE PENGURANGAN PLASTIK PADA MEDIA SOSIAL MENGGUNAKAN METODE SVM," *J. Inform. Teknol. dan Sains*, vol. 7, no. 1, hal. 147–157, 2025.
- [11] R. D. Pebrianti, "ANALISIS SENTIMEN MASYARAKAT PLATFORM X TERHADAP KORUPSI PT. PERTAMINA (PERSERO) MENGGUNAKAN METODE SVM," *JITET (Jurnal Inform. dan Tek. Elektro Ter.*, vol. 13, no. 2, 2025.
- [12] F. A. Ryandi, D. Pratiwi, dan S. Sari, "Analisis Sentimen Masyarakat Terhadap Kasus Judi Online Menggunakan Data Dari Media Sosial X Pendekatan Naive Bayes Dan SVM," *J. Sains dan Teknol.*, vol. 7, no. 1, hal. 1–6, 2025, doi: 10.47080/simika.v8i1.3624.
- [13] F. Fathoni, A. Ibrahim, F. Rizka Mumtaz, M. Azmi Zaky, M. Jodi Pratama, dan I. Akbar Kurniawan, "Analisis Sentimen Public Twitter Terhadap Kebijakan Pemerintah Menggunakan Metode SVM (Studi Kasus : RUU TNI)," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 4, hal. 6322–6329, 2025, doi: 10.36040/jati.v9i4.14036.
- [14] A. Farhan, I. Ridat, dan A. R. Isnain, "PEMODELAN TOPIK DENGAN LDA UNTUK MEMAHAMI PERSEPSI PUBLIK TERHADAP APPLE VISION PRO MELALUI ANALISIS BERITA TWITTER," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 10, no. 4, hal. 3233–3246, 2025.
- [15] M. Ulfa Yanuar dan W. Wibowo, "Topic Modeling and Sentiment Analysis on Trans Jatim Application User Reviews," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 6, no. 1, hal. 156–166, 2026.
- [16] N. N. A. Aryanti dan O. Suria, "ANALISIS SENTIMEN TERHADAP PEMUTUSAN HUBUNGAN KERJA DI INDONESIA : KOMPARASI INDOBERT DENGAN SVM, RANDOM FOREST, DAN DECISION TREE DENGAN OPTIMASI TF - IDF," *RABIT J. Teknol. dan Sist. Inf. Univrab*, vol. 10, no. 2, hal. 1158–1176, 2025.
- [17] M. H. Tinambunan, A. Hasibuan, S. Wahyuni, dan A. S. Wibowo, "Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Fasilitas Pada Ftik Universitas Dharmawangsa Medan Dengan Algoritma Naive Bayes," *Bisnis-Net J. Ekon. dan Bisnis*, vol. 6, no. 1, hal. 208–215, 2023, doi: 10.46576/bn.v6i1.3356.