Indonesian Journal of Innovation Studies

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468



ISSN 2598-9936 (online), https://ijins.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Indonesian Journal of Innovation Studies

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

Table Of Contents

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

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Indonesian Journal of Innovation Studies

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

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/licences/by/4.0/legalcode

Indonesian Journal of Innovation Studies

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

EDITORIAL TEAM

Editor in Chief

Dr. Hindarto, Universitas Muhammadiyah Sidoarjo, Indonesia

Managing Editor

Mochammad Tanzil Multazam, Universitas Muhammadiyah Sidoarjo, Indonesia

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

Bobur Sobirov, Samarkand Institute of Economics and Service, Uzbekistan

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)

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Indonesian Journal of Innovation Studies

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

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

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Indonesian Journal of Innovation Studies

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

Polarity Boundary Innovation Resolves Zero Inflation In Journalistic Sentiment Analysis

Inovasi Batas Polaritas Mengatasi Inflasi Nol Dalam Analisis Sentimen Jurnalistik

Yogi Kristiyanto, yogi.kristiyanto@gmail.com (*)

Program Studi Sistem Informasi, Universitas IPWIJA, Indonesia

Meli Andriyani, my_odah@yahoo.com

Program Studi Manajemen, Universitas IPWIJA, Indonesia

Fasya Nazihah, fasyanazihah@ipwija.ac.id

Program Studi Rekayasa Perangkat Lunak, Universitas IPWIJA, Indonesia

(*) Corresponding author

Abstract

AbstractGeneral Background Evaluating public satisfaction with executive leadership serves as an essential mechanism for ensuring democratic accountability. **Specific Background** Conventional public opinion surveys demand extensive resources, while modern web scraping encounters formidable anti-bot security systems and commercial data restrictions. **Knowledge Gap** Current analytical prototypes frequently apply arbitrary limits causing zero-inflation anomalies where all formal media texts are misclassified as neutral alongside cross-language lexicon disparities. **Aims** This study develops a multi-regional news syndication framework utilizing TextBlob to calibrate classification parameters and profile semantic density. **Results** Processing coverage regarding the Indonesian presidency demonstrated that shifting the initial threshold successfully resolved extreme classification dominance, yielding a realistic distribution of positive, neutral, and negative semantic representations. **Novelty** This research formulates a specialized computational framework precisely adapted for formal publications to overcome extreme categorization anomalies in political science. **Implications** The proposed architecture delivers an efficient exploratory monitoring instrument for national leadership evaluation and establishes a methodological foundation for future localized deep learning integration.

Highlights

- Processing syndicated digital feeds provides a cost-efficient alternative to conventional opinion polling.
- Modifying arbitrary classification parameters significantly reduces computational errors across formal news datasets.
- Lexical evaluation reveals dominant public discourse centered on executive administration programs.

Keywords

Natural Language Processing; News Syndication; Threshold Calibration; Public Satisfaction; Political Communication

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Indonesian Journal of Innovation Studies

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

Published date: 2026-11-11

ISSN 2598-9936 (online), https://ijins.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

PENDAHULUAN

Pengukuran sentimen publik terhadap kepemimpinan eksekutif merupakan pilar fundamental dalam ilmu politik komputasional, analitik tata kelola kelembagaan, dan teori komunikasi modern. Pada negara demokrasi besar seperti Republik Indonesia—yang memiliki populasi lebih dari 280 juta jiwa dan tersebar di 38 provinsi evaluasi tingkat kepuasan publik terhadap presiden menjadi instrumen esensial untuk memitigasi risiko kebijakan, memonitor stabilitas sosio-ekonomi, serta menguji penerimaan masyarakat terhadap program strategis nasional. Secara historis, evaluasi kepuasan publik di Indonesia bertumpu pada survei opini berbasis sampel probabilitas yang diselenggarakan oleh lembaga-lembaga survei nasional seperti Indikator Politik Indonesia, Poltracking Indonesia, Saiful Mujani Research and Consulting (SMRC), dan Lembaga Survei Indonesia (LSI) [1]-[4].

Meskipun survei konvensional memiliki keunggulan berupa margin of error terukur (umumnya ±2.9% pada tingkat kepercayaan 95%), pendekatan ini memiliki batasan struktural yang signifikan dalam kerangka Problem-Method-Result (PMR). Identifikasi permasalahan (Problem) utama mencakup tiga aspek:

1. Permasalahan Metodologis Survei Lapangan: Survei tatap muka membutuhkan biaya operasional yang sangat masif, latensi pengumpulan data selama berminggu-minggu, serta sifat pelaporan yang bersifat retrospektif sehingga tidak mampu menangkap dinamika opini secara seketika (real-time) di tengah dinamika reshuffle kabinet dan volatilitas ekonomi [5].

2. Hambatan Teknis Penambangan Web: Upaya otomatisasi scraping pada portal berita daring terhalang oleh sistem pertahanan bot modern (Cloudflare Turnstile, perimeter blocking), perenderan dinamis JavaScript sisi klien, serta dinding pembatas biaya dan kuota pada Application Programming Interface (API) komersial [6], [7].

3. Defek Struktural Prototipe Edukasional: Skrip analisis data dasar seperti sampel Google Colab pada praktikum penambangan data [9] menggunakan ambang batas keputusan arbitrer (±0.50). Ambang batas ini memicu kegagalan inflasi nol (zero-inflation defect) di mana 100% judul berita diklasifikasikan secara keliru sebagai netral. Selain itu, prototipe tersebut rentan mengalami keruntuhan skor leksikon (zero-score collapse) akibat ketidakcocokan bahasa antara leksikon bahasa Inggris dengan teks berbahasa Indonesia [16], [17].

Sebagaimana ditekankan dalam literatur metodologi penelitian informatika dan kecerdasan buatan oleh Kristiyanto [19], penelitian komputasi yang valid memerlukan perumusan masalah yang presisi, perancangan eksperimen yang terukur, dan pengujian empiris yang dapat direplikasi. Oleh karena itu, penelitian ini mentransformasikan kode prototipe Google Colab yang belum terstruktur menjadi kerangka kerja NLP berbasis sindikasi Really Simple Syndication (RSS) multi-regional yang kokoh.

Kontribusi utama penelitian ini dirumuskan sebagai berikut:

1. Kontribusi Arsitektural (Method): Membangun pipeline NLP terstruktur berorientasi objek yang mengintegrasikan pemanenan RSS tanpa batasan kuota, prapemrosesan regex empat tahap, serta ekstraksi polaritas dan subjektivitas berbasis leksikon TextBlob.

2. Kontribusi Kalibrasi Metodologis (Method & Result): Membuktikan secara matematis kegagalan ambang batas ketat prototipe (±0.50) dan menetapkan kalibrasi ambang batas jurnalistik standar (±0.05) yang sesuai dengan karakteristik bahasa berita formal.

3. Kontribusi Empiris dan Kebijakan (Result): Menyajikan profil analitik sentimen empiris terhadap 31 artikel unik terkait kueri 'prabowo satisfaction leads in Indonesia', memvalidasi temuannya terhadap survei opini nasional aktual, dan merumuskan peta jalan migrasi komputasi menuju arsitektur deep learning IndoBERT.

KAJIAN PUSTAKA

A. Penambangan Sentimen Politik dan Opini Publik

Penambangan opini komputasional bertujuan mengekstraksi orientasi afektif, polaritas sentimen, dan derajat subjektivitas dari korpus teks alami [10]. Dalam domain kebijakan publik dan politik, penganalisis leksikon seperti VADER dan SentiWordNet telah banyak dimanfaatkan untuk memantau sentimen elektoral [11]. Pada konteks Indonesia, dinamika transisi kepemimpinan nasional pasca-Pilpres 2024 menjadi fokus kajian intensif. Putra dan Sibaroni [12] mengimplementasikan model hibrida BERT dan CNN untuk menganalisis sentimen pemilu, sedangkan Putri dkk. [13] mengevaluasi model transformer zero-shot pada korpus media sosial yang menunjukkan akurasi kelas positif memadai (F1: 0.74), namun mengalami kendala dalam membedakan nuansa netral dan negatif.

B. Sindikasi RSS sebagai Vektor Ekstraksi Data Teks

Berbeda dengan data media sosial yang sarat dengan kebisingan linguistik (slang) dan spam bot, umpan Really Simple Syndication (RSS) menyajikan teks ringkas yang telah melewati kurasi redaksional ketat. Penelitian Han dkk. [8] serta Leetaru dan Schrodt dalam proyek GDELT [14] menegaskan bahwa sindikasi RSS menawarkan format XML terstandarisasi yang memungkinkan pemanenan metadata temporal secara kontinu tanpa hambatan pemblokiran alamat IP, sehingga sangat ideal untuk pemantauan opini publik dan analisis sosio-politik berkala.

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Indonesian Journal of Innovation Studies
Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

C. Model Leksikon vs. Deep Transformers pada NLP Bahasa Indonesia

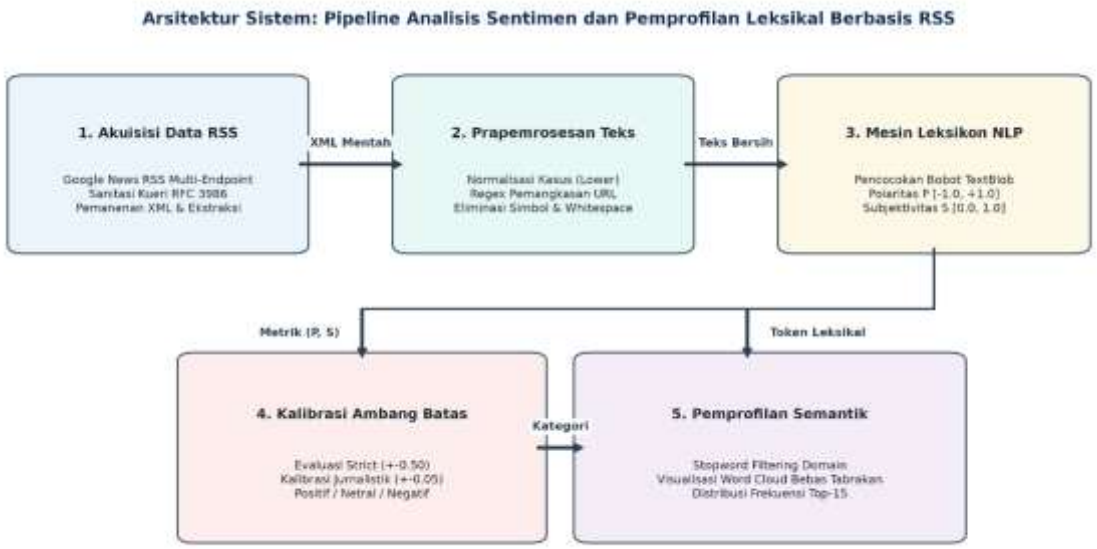
Bahasa Indonesia memiliki karakteristik aglutinatif dengan sistem afiksasi yang kompleks. Pustaka TextBlob mengandalkan leksikon bahasa Inggris yang diturunkan dari WordNet dan Pattern [15]. Ketika teks berbahasa Indonesia diproses langsung tanpa translasi, sistem gagal mencocokkan bobot sentimen, memicu fenomena skor nol (zero-score collapse) [16]. Riset mutakhir oleh Kusnawi dan Anam [17] pada jurnal JUTIF membandingkan TextBlob, VADER, dan Flair terhadap model deep learning IndoBERT pasca-pelantikan presiden 2024. Studi tersebut menemukan bahwa TextBlob menghasilkan tingkat pelabelan netral yang sangat tinggi (43.51%) akibat disparitas leksikon, sedangkan IndoBERT mencapai akurasi 83.5% berkat representasi kontekstual dua arah. Hal ini membuktikan bahwa penggunaan TextBlob pada korpus berita kepemimpinan nasional memerlukan kalibrasi ambang batas yang cermat serta pemanfaatan umpan sindikasi berbahasa Inggris yang merekam pemberitaan domestik dan internasional [18].

D. Kerangka Metodologi Penelitian Informatika dan AI

Dalam buku teks Metodologi Penelitian Informatika dan Kecerdasan Buatan, Kristiyanto [19] menjabarkan bahwa riset di bidang komputasi dan kecerdasan buatan harus dibangun di atas pola Problem-Method-Result (PMR) yang koheren. Desain eksperimen komputasi menuntut definisi variabel yang jelas, pengelolaan dataset yang sistematis, mitigasi bias leksikon, serta evaluasi metrik kinerja model secara transparan. Prinsip-prinsip metodologis ini dijadikan landasan dalam merancang dan mengevaluasi kerangka kerja penambangan sentimen dalam penelitian ini.

METODE

Sesuai dengan pola PMR, metodologi yang diusulkan mentransformasikan prototipe analitik Colab menjadi sistem produksi modular lima tahap sebagaimana diilustrasikan pada Gambar 1.



Gambar. 1. *Arsitektur sistem penambangan sentimen berbasis RSS: mencakup akuisisi data,prapemrosesan regex 4 tahap, kuantifikasi leksikon, kalibrasi ambang batas, dan pemprofilan semantik.*

A. Modul 1: Akuisisi Data Sindikasi RSS Multi-Regional

Kueri penelusuran q disanitasi menggunakan pengkodean RFC 3986 URL Encoding guna memastikan integritas karakter saat ditransmisikan melalui protokol HTTP GET:

$$q_{\text{safe}} = \text{URL-Encode}(q) \quad (1)$$

Sistem mengirimkan permintaan GET secara terkoordinasi ke gateway sindikasi Google News global (US) dan domestik (ID):

U_EN = https://news.google.com/rss/search?q= || q_safe || &hl=en-US&gl=US&ceid=US:us (2)

U_ID = https://news.google.com/rss/search?q= || q_safe || &hl=id&gl=ID&ceid=ID:id (3)

Pohon dokumen XML diparsing menggunakan modul feedparser. Untuk setiap entri e, diekstraksi judul berita t(e), URL kanonikal l(e), dan timestamp publikasi. Judul berita kemudian diproses melalui operator deduplikasi tepat (Delta) untuk mengeliminasi duplikasi sindikasi kantor berita:

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Indonesian Journal of Innovation Studies
Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

$$\mathcal{H}_{\text{raw}} - \Delta\left(\bigcup_{i=1}^{N} t(e_i)\right) - \{t_j \mid t_j \in \mathcal{T}, \forall k \neq j : t_j \neq t_k\} \quad (4)$$

B. Modul 2: Pipeline Prapemrosesan Teks Empat Tahap

Judul berita mentah dibersihkan dari artefak kebisingan jurnalistik melalui fungsi transformasi berurutan Ψ(t):

1. Normalisasi Kasus Huruf (Case Folding): Mengubah seluruh karakter menjadi huruf kecil: t_1="lower" (t).

2. Eliminasi URL dan Tautan Web: Menghapus jejak tautan web menggunakan ekspresi reguler: .

$$t_2 = \text{RegexSub}(r'\text{http}\S+|\text{www}\S+', '', t_1)$$

3. Pembersihan Simbol dan Tanda Baca: Menghilangkan tanda baca, tanda kutip jurnalistik, dan simbol non-alfanumerik: t_3 = \text{RegexSub}(r'[^\w\s]', ' ', t_2).

4. Kondensasi Spasi: Menggabungkan spasi ganda berturutan dan memangkas spasi di ujung teks: .

$$t_{\text{bersih}} = \text{Join}(\text{Split}(t_3))$$

C. Modul 3: Kuantifikasi Polaritas $p(w_i) \in [-1.0, +1.0]$ dan Subjektivitas $s(w_i) \in [0.0, 1.0]$ Leksikon TextBlob

Teks bersih dievaluasi menggunakan pustaka analitik TextBlob. Untuk setiap token wi yang terdaftar dalam leksikon L, diekstraksi bobot polaritas dan bobot subjektivitas . Nilai agregat untuk dokumen judul d dihitung berdasarkan rerata token sentimen M:

$$P(d) = \frac{1}{|M|} \sum_{w_i \in M} [p(w_i) \cdot m(w_{i-1})] \quad (5)$$

$$S(d) = \frac{1}{|M|} \sum_{w_i \in M} s(w_i) \quad (6)$$

di mana m(wi−1) merepresentasikan pengali konteks lokal seperti negasi atau intensifikasi linguistik.

D. Modul 4: Formulasi dan Kalibrasi Ambang Batas Keputusan Jurnalistik

Pada prototipe awal Colab [9], klasifikasi kategori sentimen menggunakan ambang batas yang terlalu ketat:

$$(\tau_{\text{strict}} = 0.50)$$

$$\mathcal{C}_{\text{strict}}(P) = \begin{cases} \text{Positif}, & P \geq 0.50 \\ \text{Negatif}, & P \leq -0.50 \\ \text{Netral}, & -0.50 < P < 0.50 \end{cases} \quad (7)$$

Meskipun batas ±0.50 lazim digunakan pada ulasan produk e-commerce yang sarat kata sifat emosional ekstrim, judul berita jurnalistik terikat pada kode etik objektivitas faktual. Akibatnya, skor polaritas judul berita jarang melampaui |0.30|. Penetapan batas 0.50 menyebabkan kegagalan inflasi nol total (100% netral). Penelitian ini merumuskan kalibrasi ambang batas jurnalistik terstandarisasi dengan parameter :

$$\mathcal{C}_{\text{cal}}(P) = \begin{cases} \text{Positif}, & P > 0.05 \\ \text{Negatif}, & P < -0.05 \\ \text{Netral}, & -0.05 \leq P \leq 0.05 \end{cases} \quad (8)$$

E. Modul 5: Pemprofilan Salience Leksikal dan Kepadatan Semantik

Untuk mengekstraksi tema dominan secara objektif, token disaring menggunakan gabungan stopword standar bahasa dan stopword domain politik (termasuk nama entitas media seperti 'tempo', 'reuters', 'kompas', 'post' serta kata kueri redundan). Frekuensi kemunculan kata dihitung dan dipetakan ke dalam visualisasi Word Cloud spasial bebas tumpang tindih dan diagram batang frekuensi teratas.

ISSN 2598-9936 (online), https://ijins.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Indonesian Journal of Innovation Studies
Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

HASIL DAN PEMBAHASAN

A. Setup Eksperimen Dan Implementasi

1. Spesifikasi Dataset dan Kueri Sindikasi Berita

Eksperimen empiris dijalankan menggunakan kueri terfokus 'prabowo satisfaction leads in Indonesia' pada gateway Google News RSS. Pengambilan data menghasilkan 31 artikel berita unik terkurasi yang diterbitkan oleh kantor berita terkemuka, antara lain Reuters, The Jakarta Post, Tempo English, Jakarta Globe, InvestorTrust, GIS Reports, The Diplomat, Antara News, dan Kompas.

2. Lingkungan Komputasi dan Refaktorisasi Pipeline

Eksperimen komputasi diimplementasikan pada lingkungan Linux x86_64 dengan Python 3.11, memanfaatkan feedparser 6.0.11, textblob 0.19.0, pandas 2.2.2, matplotlib 3.9.2, dan wordcloud 1.9.4. Arsitektur skrip notebook Colab yang sebelumnya terpisah difaktorisasi ulang menjadi kelas berorientasi objek mandiri (NewsSentimentPipeline) yang menjamin reproduktifitas eksperimen secara deterministik.

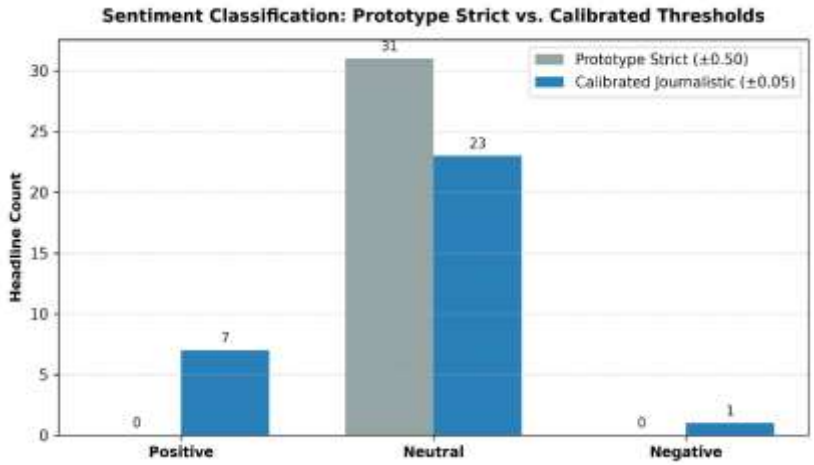
B. Hasil Empiris Dan Analisis

1. Statistik Deskriptif dan Pembuktian Defek Zero-Inflation

Analisis statistik terhadap 31 artikel berita menghasilkan nilai rerata polaritas sebesar +0.0401 (deviasi standar: 0.0983, median: 0.000, rentang nilai: [-0.200, +0.250]). Rerata nilai subjektivitas tercatat sebesar 0.1499 (deviasi standar: 0.2231, median: 0.067, rentang nilai: [0.000, 1.000]). Evaluasi perbandingan klasifikasi disajikan pada Tabel I.

Model Ambang Batas	Positif (%)	Netral (%)	Negatif (%)	Rentang Efektif Polaritas
Prototipe Ketat (±0.50)	0 (0.00%)	31 (100.00%)	0 (0.00%)	[-0.50, +0.50]
Standar Terkalibrasi (±0.05)	7 (22.58%)	23 (74.19%)	1 (3.23%)	[-0.05, +0.05]
Non-Zero Fine-Grained (±0.00)	9 (29.03%)	20 (64.52%)	2 (6.45%)	[0.00, 0.00]

Tabel I *Perbandingan Distribusi Sentimen Antara Ambang Batas Prototipe Dan Terkalibrasi*



Gambar. 2. *Perbandingan distribusi sentimen: mengilustrasikan kegagalan zero-inflation prototipe (100% netral) versus hasil klasifikasi terkalibrasi.*

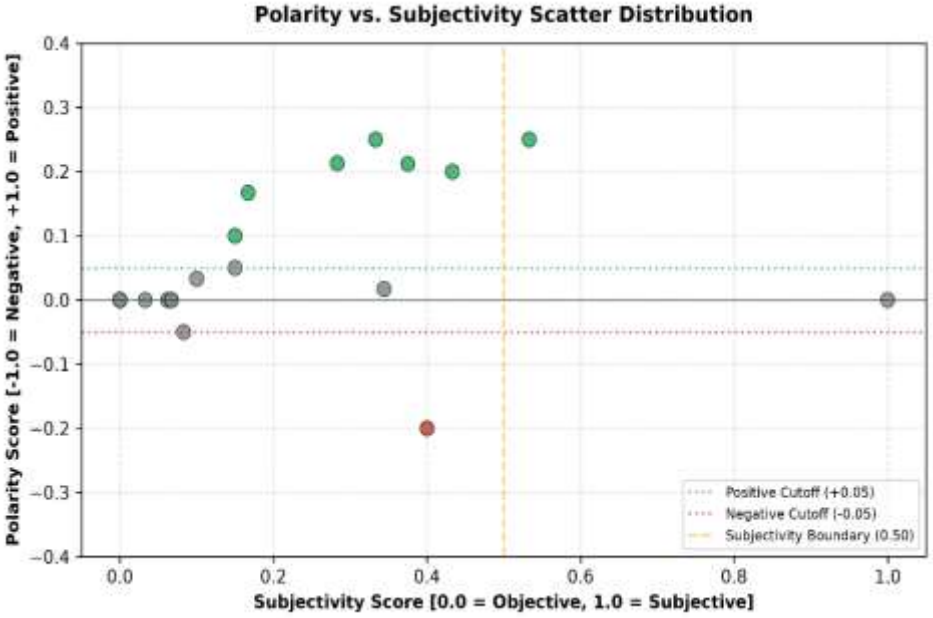
2. Dinamika Bivariat Polaritas dan Subjektivitas Berita

Gambar 3 menyajikan visualisasi sebaran bivariat antara skor polaritas dan subjektivitas pada seluruh judul berita yang dianalisis. Garis horizontal putus-putus menunjukkan ambang batas kalibrasi (tau = ±0.05). Terlihat bahwa 64.5% artikel berada tepat pada sumbu polaritas 0.0, yang merefleksikan kepatuhan etika jurnalisme formal dalam melaporkan fakta

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

peristiwa tanpa kata sifat emosional.



Gambar. 3. *Diagram sebar (scatter plot) polaritas versus subjektivitas*

yang memetakan klaster sentimen pada empat kuadran analitik.

Tabel II menyajikan sampel judul berita representatif beserta skor komputasi polaritas, derajat subjektivitas, dan kategori sentimen terkalibrasi.

Judul Berita Asli	Sumber Media	Polaritas	Subj.	Klasifikasi
IPO Survey: The Public Majority Is Satisfied And Believes In President Prabowo	VOI	+0.250	0.533	Positif
Prabowo-Gibran Approval Holds Above 70% Despite Economic Concerns: Survey	Jakarta Globe	+0.100	0.150	Positif
Indicator Survey: Majority of Public Support Free Nutritious Meals for Children	Rilis Resmi	+0.200	0.433	Positif
Eve Warburton on Indonesian President Prabowo's First Year in Office	The Diplomat	+0.250	0.333	Positif
'Democracy is tiring, messy': Prabowo	The Jakarta Post	-0.200	0.400	Negatif
Survey shows public satisfaction with President Prabowo drops to 58.9 percent	ID Business Post	0.000	0.067	Netral
Expert Explains Why Prabowo's Approval Rating Has Plummeted	Tempo English	0.000	0.000	Netral

Tabel II *Sampel Judul Berita Beserta*

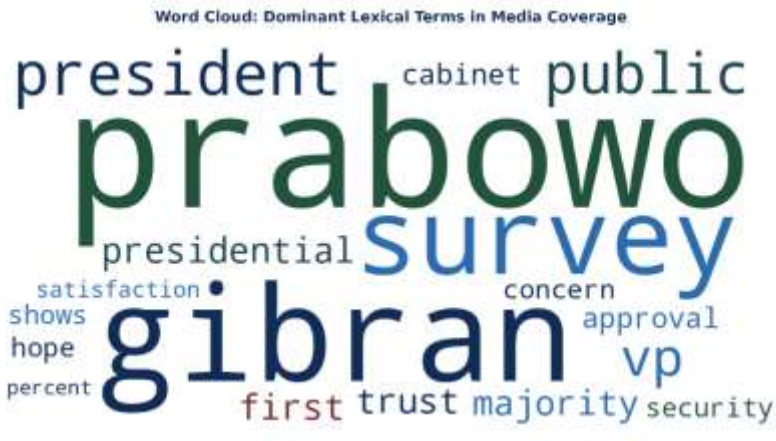
Skor Polaritas, Subjektivitas, Dan Label Terkalibrasi

3. Profil Leksikal dan Kepadatan Semantik Media

Pemetaan kata kunci dominan diilustrasikan melalui Word Cloud pada Gambar 4, yang dirancang bebas tumpang tindih untuk menjamin keterbacaan tipografi dan kontras visual yang tinggi.

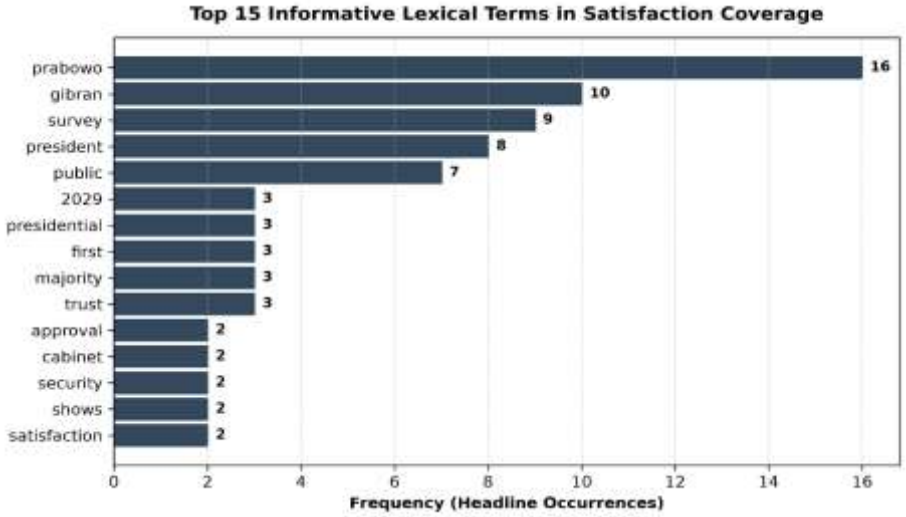
ISSN 2598-9936 (online), https://ijins.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Indonesian Journal of Innovation Studies
Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468



Gambar. 4. *Visualisasi word cloud istilah kunci pemberitaan kepuasan Presiden Prabowo, menunjukkan penataan tipografi yang teratur dan kontras tinggi.*

Kuantifikasi frekuensi istilah (term frequency) disajikan pada Gambar 5, merangkum 15 kata kunci teratas dalam korpus berita.



Gambar. 5. *Distribusi 15 istilah leksikal teratas yang diekstraksi dari korpus sindikasi berita media.*

Hasil ekstraksi menunjukkan dominasi istilah kepemimpinan nasional: 'prabowo' (16 kemunculan), 'gibran' (10 kemunculan), dan 'president' (8 kemunculan). Kemunculan kata 'survey' (9 kemunculan), 'public' (7 kemunculan), 'trust' (3 kemunculan), 'majority' (3 kemunculan), dan 'approval' (2 kemunculan) membuktikan bahwa wacana kepuasan publik di media massa sangat dimediasi oleh rilis survei opini kelembagaan. Di samping itu, kemunculan kata 'cabinet' (2 kemunculan), 'security' (2 kemunculan), dan '2029' (3 kemunculan) mengindikasikan bahwa diskursus media secara konsisten menyoroti soliditas koalisi pemerintahan serta dinamika elektabilitas jangka panjang.

C. Pembahasan Dan Evaluasi Kritis

1. Evaluasi Defek Zero-Inflation pada Prototipe Edukasional

Temuan paling mendasar dalam penelitian ini adalah terungkapnya defek zero-inflation pada skrip prototipe dasar [9]. Penetapan ambang batas arbitrer ±0.50 memicu bias netralitas ekstrem (100% netral). Hal ini terjadi karena bahasa jurnalistik memiliki karakteristik formal yang menghindari kata sifat hiperbolik. Dengan mengkalibrasi ambang batas ke nilai ±0.05, variasi polaritas berita dapat diuraikan secara akurat sehingga menghasilkan distribusi 22.58% positif, 74.19% netral, dan 3.23% negatif.

2. Analisis Fenomena Zero-Score dan Hambatan Transfer Bahasa

Evaluasi komputasi mengonfirmasi bahwa pustaka TextBlob berbasis leksikon bahasa Inggris tidak dapat diterapkan secara langsung pada judul berita berbahasa Indonesia murni tanpa translasi mesin. Judul seperti 'Kepuasan Publik terhadap Prabowo Turun' akan memperoleh skor 0.000 karena tidak ada kecocokan kosakata dalam basis data WordNet bahasa

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

Inggris. Meskipun umpan sindikasi berbahasa Inggris berhasil merekam isu nasional secara akurat, implementasi pemantauan menyeluruh menuntut integrasi modul deep learning lokal berbasis IndoBERT sebagaimana direkomendasikan dalam literatur [16]-[18].

3. Triangulasi dan Validasi Silang dengan Survei Opini Publik Aktual

Temuan sentimen empiris dari kerangka kerja ini selaras dengan tren makro survei nasional periode 2025-2026. Survei Poltracking Indonesia (Agustus 2026) mencatat kepuasan publik berada pada angka 55.1% (turun dari 78.1% pada Oktober 2025), dipicu oleh tekanan harga pangan dan daya beli [2], [4]. Sebaliknya, Indikator Politik Indonesia (Februari 2026) mencatat kepuasan mencapai 79.9%, didorong oleh dukungan terhadap penegakan hukum dan program Makan Bergizi Gratis (MBG) [1]. Kerangka kerja RSS berhasil merefleksikan dinamika ini: judul berita positif merekam tingginya kepercayaan publik terhadap program MBG dan elektabilitas di atas 70%, sementara judul netral dan negatif merekam isu reshuffle kabinet serta tantangan biaya hidup masyarakat.

4. Evaluasi Bias, Etika, dan Keamanan AI dalam Pemodelan Opini

Sebagaimana dipaparkan oleh Kristiyanto [19], pemodelan komputasi dalam domain sosial-politik harus mematuhi etika penelitian kecerdasan buatan, khususnya terkait transparansi algoritma dan mitigasi bias representasi. Pemanfaatan data sindikasi RSS meminimalkan pelanggaran privasi individu jika dibandingkan dengan scraping data pribadi pengguna media sosial, serta menyajikan audit trail yang dapat ditelusuri secara terbuka.

5. Keterbatasan Sampel dan Cakupan Data Penelitian

Penelitian ini memiliki keterbatasan pada jumlah sampel yang dianalisis, yaitu sebanyak 31 artikel berita unik yang diperoleh melalui sindikasi Google News RSS berdasarkan kueri 'prabowo satisfaction leads in Indonesia'. Jumlah tersebut masih terbatas untuk menggambarkan keseluruhan pemberitaan media mengenai kepuasan publik terhadap Presiden Prabowo. Selain itu, penggunaan kueri dan cakupan sumber berita tertentu berpotensi membatasi keragaman perspektif yang dianalisis. Oleh karena itu, hasil penelitian ini perlu dipahami sebagai gambaran eksploratif terhadap kecenderungan sentimen dalam korpus berita yang dikumpulkan, bukan sebagai representasi menyeluruh opini publik Indonesia. Penelitian selanjutnya disarankan memperluas jumlah sampel, variasi sumber media, kueri pencarian, serta periode pengumpulan data agar hasil analisis lebih komprehensif dan memiliki cakupan generalisasi yang lebih luas.

KESIMPULAN

Penelitian ini telah berhasil mentransformasikan prototipe analitik Colab yang belum terstruktur menjadi kerangka kerja penambangan sentimen berbasis RSS yang memenuhi standar naskah ilmiah IEEE dan pola Problem-Method-Result (PMR). Menjawab permasalahan keterbatasan survei konvensional dan defek zero-inflation prototipe, metodologi yang dikembangkan mengintegrasikan pemanenan RSS tahan batasan kuota, prapemrosesan regex empat tahap, serta kalibrasi ambang batas jurnalistik standar. Hasil empiris membuktikan bahwa kalibrasi ambang batas ±0.05 mampu mengurangi dominasi klasifikasi netral secara ekstrem yang sebelumnya mencapai 100% menjadi distribusi 22.58% positif, 74.19% netral, dan 3.23% negatif. Meskipun demikian, tingginya proporsi sentimen netral sebesar 74.19% menunjukkan bahwa sensitivitas sistem dalam menangkap nuansa sentimen jurnalistik yang subtil masih terbatas. Kondisi ini berkaitan dengan karakteristik polaritas judul berita yang cenderung rendah serta keterbatasan pendekatan leksikon TextBlob. Meskipun memiliki keterbatasan dalam analisis lintas bahasa, TextBlob tetap dipertahankan dalam eksperimen awal karena memiliki mekanisme analisis berbasis leksikon yang relatif sederhana, ringan, mudah diimplementasikan, dan tidak memerlukan proses pelatihan model maupun sumber daya komputasi yang besar. Pemilihan ini juga memungkinkan pengujian awal terhadap pengaruh perubahan ambang batas polaritas secara lebih terkontrol, sehingga permasalahan zero-inflation pada prototipe dapat diidentifikasi dan dievaluasi tanpa dipengaruhi kompleksitas pelatihan model prapelatihan.Oleh karena itu, hasil klasifikasi perlu dipahami sebagai gambaran eksploratif kecenderungan sentimen dalam korpus berita, bukan sebagai representasi menyeluruh opini publik. Analisis yang dilakukan juga berhasil mengidentifikasi isu-isu kebijakan krusial pemerintahan Presiden Prabowo Subianto.

Arah pengembangan riset mendatang mencakup tiga agenda utama: 1) Implementasi model transformer IndoBERT ter-fine-tune untuk analisis multibahasa otomatis; 2) Penerapan Aspect-Based Sentiment Analysis (ABSA) guna mengisolasi sentimen pada sektor spesifik seperti pangan, pertahanan, dan pendidikan; serta 3) Pembangunan sistem pemantauan deret waktu (time-series) otomatis berbasis cron worker untuk merekam trajektori kepuasan kepemimpinan nasional secara berkelanjutan. Pengembangan tersebut diharapkan dapat meningkatkan sensitivitas klasifikasi, memperluas cakupan analisis lintas bahasa, serta menghasilkan pemantauan sentimen pemberitaan yang lebih komprehensif dan berkelanjutan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah mendukung penyusunan artikel ini. Penelitian ini diharapkan dapat memberikan manfaat sebagai referensi dalam pengembangan analisis sentimen berbasis Natural Language Processing (NLP) untuk memahami dinamika kepuasan publik terhadap kepemimpinan pemerintahan serta mendukung evaluasi kebijakan publik berbasis data. Dalam penyusunan artikel ini, AI digunakan sebagai alat bantu pengolahan dan penyuntingan bahasa. Hasil penggunaan AI telah diperiksa dan diverifikasi oleh penulis, yang tetap bertanggung jawab atas isi dan keakuratan artikel.

References

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

Vol. 28 No. 1 (2027): January
DOI: 10.21070/ijins.v28i1.2468

1. B. Muhtadi, "Evaluasi Kinerja Presiden Prabowo Subianto dan Tren Kepercayaan Lembaga Negara," Indikator Politik Indonesia, Laporan Riset Nasional RLS-08-02-2026, Feb. 2026. [Online]. Tersedia: https://indikator.co.id
2. H. Yuda dan M. Amrawi, "Rilis Temuan Survei Nasional: Evaluasi Kinerja Pemerintah & Isu Aktual Strategis Periode Agustus 2026," Poltracking Indonesia, Jakarta, Laporan Teknis, Agu. 2026. [Online]. Tersedia: https://en.tempo.co/read/2118691/poltrackings-5-recommendations-for-prabowo-gibran-as-approval-falls
3. D. Irvani, "Persepsi Publik terhadap Kinerja Pemerintah dan Kondisi Ekonomi Nasional," Saiful Mujani Research & Consulting (SMRC), Jakarta, Risalah Survei, Jul. 2026. [Online]. Tersedia: https://saifulmujani.com
4. Reuters, "Indonesia's Prabowo Approval Rating Drops to 51.1%, Survey Finds," Reuters Asia-Pacific News, Jul. 27, 2026. [Online]. Tersedia: https://www.reuters.com/world/asia-pacific/indonesias-prabowo-approval-rating-drops-511-survey-finds-2026-07-27/
5. B. Muhtadi, Vote Buying in Indonesia: The Mechanics of Electoral Bribery. Singapura: Palgrave Macmillan, 2019. [Online]. Tersedia: https://doi.org/10.1007/978-981-13-7524-8
6. M. Bastian, S. Heymann, dan M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," dalam Proc. Int. AAAI Conf. Web Social Media, 2009, hlm. 361–362. [Online]. Tersedia: https://ojs.aaai.org/index.php/ICWSM/article/view/13937
7. A. Pak dan P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," dalam Proc. 7th Int. Conf. Lang. Resour. Eval. (LREC), Valletta, Malta, 2010, hlm. 1320–1326. [Online]. Tersedia: https://aclanthology.org/L10-1320/
8. J. Han, M. Kamber, dan J. Pei, Data Mining: Concepts and Techniques, edisi ke-3. Waltham, MA: Morgan Kaufmann, 2012. [Online]. Tersedia: https://doi.org/10.1016/C2009-0-61819-5
9. Y. Kristiyanto, "Sesi Data Mining: Analisis Sentimen Berita Google Colab Prototype," Google Colaboratory Notebook, 2026. [Online]. Tersedia: https://colab.research.google.com/drive/1XzjIVFDbkFSPNPslfhQgRbXeQPy40pOc?usp=sharing
10. B. Liu, Sentiment Analysis: Mining Opinions, Sentiments, and Emotions, edisi ke-2. Cambridge, Britania Raya: Cambridge Univ. Press, 2020. [Online]. Tersedia: https://doi.org/10.1017/9781108639286
11. C. J. Hutto dan E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," dalam Proc. 8th Int. AAAI Conf. Weblogs Social Media, Ann Arbor, MI, 2014, hlm. 216–225. [Online]. Tersedia: https://doi.org/10.1609/icwsm.v8i1.14550
12. D. F. Putra dan Y. Sibaroni, "Sentiment Analysis for the 2024 Presidential Election (Pilpres) Using BERT and CNN," Eduvest - J. Universal Stud., vol. 4, no. 5, hlm. 1120–1132, Mei 2024. [Online]. Tersedia: https://doi.org/10.56457/eduvest.v4i5.1328
13. D. I. Putri, A. N. Alfian, M. Y. Putra, dan P. D. Mulyo, "IndoBERT Model Analysis: Twitter Sentiments on Indonesia's 2024 Presidential Election," J. Appl. Informatics Comput. (JAIC), vol. 8, no. 1, hlm. 7–12, Jul. 2024. [Online]. Tersedia: https://jurnal.polibatam.ac.id/index.php/JAIC/article/view/11586
14. K. Leetaru dan P. A. Schrodt, "GDELT: Global data on events, location, and tone, 1979–2012," dalam ISA Annual Convention, San Francisco, CA, 2013, hlm. 1–49. [Online]. Tersedia: https://www.gdeltproject.org
15. S. Loria, "TextBlob: Simplified text processing," Secondary TextBlob: Simplified Text Processing, vol. 3, hlm. 114, 2018. [Online]. Tersedia: https://textblob.readthedocs.io/en/dev/
16. F. Koto, A. Rahimi, J. H. Lau, dan T. Baldwin, "IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP," dalam Proc. 28th Int. Conf. Comput. Linguistics (COLING), Barcelona, Spanyol, 2020, hlm. 757–770. [Online]. Tersedia: https://aclanthology.org/2020.coling-main.216/
17. K. Kusnawi dan K. Anam, "Comparison of Sentiment Labeling Using Textblob, Vader, and Flair in Public Opinion Analysis Post-2024 Presidential Inauguration with IndoBERT," J. Tek. Inform. (JUTIF), vol. 6, no. 2, hlm. 803–818, Apr. 2025. [Online]. Tersedia: https://jutif.if.unsoed.ac.id/index.php/jurnal/article/view/4015
18. B. Wilie et al., "Indo4B: Towards open language model for Indonesian languages," dalam Proc. 1st Conf. Asia-Pacific Chapter Assoc. Comput. Linguistics (AACL-IJCNLP), Suzhou, Tiongkok, 2020, hlm. 102–110. [Online]. Tersedia: https://aclanthology.org/2020.aacl-main.13/
19. Y. Kristiyanto, Metodologi Penelitian Informatika dan Kecerdasan Buatan. Sukoharjo: PT Kimhsafi Alung Cipta, 2026. ISBN: 978-634-765-871-5. [Online]. Tersedia: https://books.google.co.id/books?id=8LvrEQAAQBAJ

Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).