

Five-Class Entity-Aware Sentiment Classification on Indonesian Environmental Discourse: KNN vs. SVM under Class Imbalance

Nandana Hadayana Tyovan¹, Raudya Lazzuard Zakkawali², Farrel Hadi Dewanto³,
Haniyah Maharani⁴, Novita Br Ginting⁵

^{1,2,3,4,5}Program Studi Sistem Informasi, Fakultas Teknik dan Sains, Universitas Ibn Khaldun Bogor, Indonesia
email: nandana.tyovan@uika-bogor.ac.id

Receive: 18 September 2026

Accepted: 30 September 2026

Abstract

Social media platforms have reshaped public discourse on environmental issues, yet Indonesian sentiment analysis has largely relied on binary or topic-centric labeling that does not distinguish who sentiment is directed at. This study analyzed Twitter/X sentiment toward two distinct target entities: eco-activists—represented by Pandawara Group, a grassroots community from Bandung—and the Indonesian government. A five-class entity-differentiated labeling scheme was devised, encoding four directional classes (positive/negative toward government; positive/negative toward activists) and one Neutral class. A dataset of 378 labeled tweets was collected via seven keywords across three iterative scraping sessions, covering posts from January 1, 2025 to May 3, 2026, with a retention rate of 47.3% after multi-stage filtering. TF-IDF feature extraction was applied to K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) classifiers, evaluated through Stratified 5-Fold Cross-Validation. KNN outperformed SVM, reaching an accuracy of 59.79% (weighted F1: 0.5651) against 49.73% (weighted F1: 0.41). Under a 22:1 class imbalance, SVM collapsed into majority-class over-prediction, while KNN maintained comparatively balanced class recognition. The sentiment distribution was consistent with two theoretical constructs—Institutional Skepticism, suggested by 41.3% negative sentiment toward government, and Parasocial Trust, suggested by only 1.9% negative sentiment toward activists—though these labels should be read as interpretive rather than confirmatory. These patterns tentatively position eco-activist communities such as Pandawara Group as a subject worth further investigation for government environmental communication, and underscore the value of entity-aware sentiment frameworks in nuanced socio-political contexts.

Keywords: sentiment analysis; entity-aware labeling; KNN; SVM; TF-IDF

Abstrak

Platform media sosial telah mengubah wacana publik mengenai isu lingkungan, namun analisis sentimen di Indonesia umumnya masih memakai pelabelan biner atau berbasis topik yang tidak membedakan kepada siapa sentimen tersebut ditujukan. Penelitian ini menganalisis sentimen Twitter/X terhadap dua entitas sasaran yang berbeda: aktivis lingkungan—direpresentasikan oleh Pandawara Group, komunitas akar rumput dari Bandung—dan pemerintah Indonesia. Skema pelabelan lima kelas berbasis entitas dirancang, mencakup empat kelas terarah (positif/negatif terhadap pemerintah; positif/negatif terhadap aktivis) dan satu kelas Netral. Sebanyak 378 tweet berlabel dikumpulkan melalui tujuh kata kunci dalam tiga sesi scraping berulang, mencakup unggahan sejak 1 Januari 2025 hingga 3 Mei 2026, dengan tingkat retensi 47,3% setelah penyaringan bertahap. Ekstraksi fitur TF-IDF diterapkan pada pengklasifikasi K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM), dievaluasi melalui Stratified 5-Fold Cross-Validation. KNN mengungguli SVM dengan akurasi 59,79% (F1 tertimbang: 0,5651) berbanding 49,73% (F1 tertimbang: 0,41). Di bawah ketidakseimbangan kelas 22:1, SVM cenderung memprediksi kelas mayoritas secara berlebihan, sementara KNN menjaga pengenalan kelas yang relatif lebih seimbang. Distribusi sentimen ini

konsisten dengan dua konstruk teoretis—Skeptisisme Institusional, yang tercermin dari 41,3% sentimen negatif terhadap pemerintah, dan Kepercayaan Parasosial, yang tercermin dari hanya 1,9% sentimen negatif terhadap aktivis—meskipun interpretasi ini bersifat sugestif, bukan konfirmatif. Temuan ini menempatkan komunitas aktivis lingkungan seperti Pandawara Group sebagai topik yang layak diteliti lebih lanjut dalam konteks komunikasi lingkungan pemerintah, sekaligus menegaskan pentingnya kerangka analisis sentimen berbasis entitas dalam konteks sosio-politik yang kompleks.

Kata Kunci: analisis sentimen; pelabelan berbasis entitas; KNN; SVM; TF-IDF

INTRODUCTION

Over the past decade, social media platforms have fundamentally altered how citizens in Indonesia engage with environmental issues. Rather than turning to official channels, many Indonesians now look to digital communities and eco-activist figures for information and inspiration on sustainability—a shift that carries meaningful implications for institutional trust (Prayogo & Adiwijaya, 2025).

Pandawara Group, a grassroots eco-activist community based in Bandung and active across Twitter/X, TikTok, and Instagram, reflects this trend. Their hands-on campaigns—cleaning rivers, beaches, and illegal dumpsites—have mobilized public engagement on a scale that official government environmental programs rarely achieve, consistent with evidence that social media influencers are now significant drivers of environmental awareness (Nazir & Wani, 2024).

At the same time, Indonesia's Ministry of Environment and Forestry (KLHK) struggles with a credibility problem. Public skepticism toward the government has grown steadily, fed by perceptions of weak regulatory enforcement, hollow green policy commitments, and inadequate responses to the country's long-running waste crisis (Hsiao & Kuipers, 2025). This points to a growing gap in trust between formal institutions and citizens.

Unlike eco-influencers whose primary role is product promotion, digital eco-activists use social media as a platform for genuine environmental advocacy. Pandawara Group exemplifies this distinction—their content-driven mobilization strategy has demonstrably raised public awareness and prompted collective environmental action across multiple regions in Indonesia (Zuliansyah & Jatmika, 2025). Community-grounded movements tend to resonate more deeply with digital audiences than corporate sustainability messaging: environmental NGOs and influencers outperform corporate brands on engagement metrics in sustainability content (Harliyono & Priskila, 2025), the perceived trustworthiness of environmental influencers directly shapes followers' environmental behavior (Nazir & Wani, 2024), and authenticity—not institutional authority—is the decisive currency of credibility in digital environmental communication (Buvár et al., 2023).

Parasocial Trust describes the trust that develops through parasocial relationships—the one-sided emotional attachments audiences form with public figures they follow on social media (Wasike, 2025; Karnjanapoomi, 2023). These parasocial connections have measurable effects on how audiences evaluate the credibility of information sources (Wasike, 2025), and relationship depth in such interactions positively predicts trust levels (Karnjanapoomi, 2023). Perceived value congruence between audience and influencer strengthens this bond further (Koay et al., 2023a), as does an influencer's willingness to share personal experience (Koay et al., 2023b)—helping explain why Indonesian audiences tend

to extend notably high trust to eco-activists like Pandawara Group (Prayogo & Adiwijaya, 2025).

Institutional Skepticism, conversely, captures citizens' disposition to question or distrust government claims and actions. Meaningful public participation mechanisms build citizen trust in government, and their absence deepens skepticism (Campbell, 2023). In the Indonesian context, fewer than 3% of local legislative candidates foregrounded environmental issues during the 2024 election cycle (Hsiao & Kuipers, 2025), and cross-nationally, social media political activity correlates negatively with trust in formal institutions while correlating positively with trust in civil society actors (Parr, 2025)—a pattern mirrored by the finding that transparency and participatory governance are the principal determinants of institutional trust (Mahmud, 2024).

Greenwashing—making environmental claims unbacked by meaningful action—further erodes stakeholder trust over time (Zioło et al., 2024), and institutional greenwashing is a primary mechanism through which public skepticism toward government environmental commitments solidifies (Santos et al., 2024). In Indonesia specifically, consumer perceptions of greenwashing weaken trust in green products and policy claims alike (Sholihah et al., 2025). Trust toward activists is not unconditional, however: commercial brand partnerships can be perceived as compromising an influencer's authenticity and backfire on credibility (Giambastiani et al., 2025), and trust depends on consistency of messaging over time (Yıldırım et al., 2024; Wang et al., 2023).

Several prior studies inform the methodological design of this research. KNN and SVM have been applied to Indonesian Twitter sentiment analysis, though typically with in-sample evaluation and without differentiating sentiment by target entity (Adharani et al., 2025). Stratified K-Fold Cross-Validation has been shown to produce more realistic performance estimates than in-sample evaluation for Indonesian election-tweet analysis (Sembiring & Dewa, 2025). SVM with TF-IDF performs competitively on Indonesian tweet classification but remains vulnerable to class imbalance without explicit mitigation (Rohmatun & Baita, 2025). The closest domain-adjacent dataset—Indonesian Twitter sentiment on biodiversity policy—uses topic-centric labeling that does not identify sentiment targets (Uliniansyah et al., 2024).

Despite this growing body of research, no prior study has applied a labeling scheme that differentiates public sentiment toward two distinct entities—eco-activists and the government—within the same environmental issue domain. Addressing this gap is the central contribution of this study, which: (1) maps the distribution of Twitter/X sentiment toward eco-activists and the government; (2) compares KNN and SVM with TF-IDF under Stratified 5-Fold Cross-Validation; and (3) interprets the findings through Institutional Skepticism and Parasocial Trust.

RESEARCH METHOD

Data Collection

Data were drawn from Twitter/X using seven keywords chosen for their coverage of Indonesian environmental public discourse: (1) Pandawara Group, (2) aktivis lingkungan, (3) KLHK sampah, (4) relawan lingkungan, (5) pencemaran lingkungan Indonesia, (6) pemerintah lingkungan hidup, and (7) eco influencer. Scraping was carried out via the

Twitter/X API using a Python script in Google Colaboratory across three iterative sessions extending temporal coverage to a final window from January 1, 2025 to May 3, 2026.

Only Indonesian-language tweets were retained. A multi-stage cleaning process—removing duplicates, emoji-only posts, and irrelevant content—reduced the initial pool of approximately 800 raw tweets to 378 usable entries (retention rate $\approx 47.3\%$).

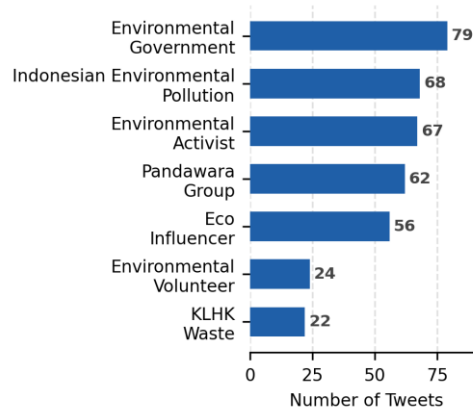


Figure 1. Tweet Distribution by Scraping Keyword (n=378)

Text Preprocessing

Preprocessing followed two stages: cleaning (stripping URLs, mentions, hashtags, numerals, and punctuation) and stopwords removal using a standard Indonesian stopwords list enriched with informal social-media terms (e.g., “gak”, “aja”, “sih”, “rt”). Stemming was deliberately omitted in this iteration.

Five-Class Labeling (Novelty)

The defining methodological contribution of this study is a labeling scheme anchored to a specific entity rather than a topic: who is a tweet directed at, and is the sentiment positive or negative toward them? This yields four directional classes—Negative/Positive toward Government, Positive/Negative toward Influencer/Activist—plus a fifth Neutral class.

Labeling was performed manually by three annotators working independently from a shared annotation guide, with final labels determined by majority vote and group discussion in case of a three-way split. Inter-rater agreement, assessed via Fleiss' Kappa, was $\kappa = 0.13$ (slight agreement per Landis & Koch, 1977), suggesting the class boundaries—particularly between Neutral and the directional classes—were not sufficiently unambiguous; this is revisited as a limitation in the Results and Discussion section.

Technical Basis: TF-IDF, KNN, and SVM

Text features were extracted using TF-IDF (max_features=5,000, ngram_range=(1,2), sublinear_tf=True), which weights terms by local frequency against corpus-wide distribution. Two classifiers were compared: KNN (k=13, cosine distance) and SVM (RBF kernel, C=1.0, gamma='scale'). The literature generally recommends a linear kernel for high-dimensional sparse features such as TF-IDF (Liu et al., 2025); the RBF choice here may partly explain SVM's weaker performance, discussed further below.

Both models were evaluated under Stratified 5-Fold Cross-Validation, preferred over in-sample evaluation because it yields honest performance estimates on unseen data for small, imbalanced datasets (Sembiring & Dewa, 2025; Zamsuri et al., 2024).

RESULTS AND DISCUSSION

Sentiment Label Distribution

The labeled dataset of 378 tweets is heavily skewed, as shown in Table 1 and Figure 2. Negative sentiment toward the government accounts for over four in ten tweets, while negative sentiment toward eco-activists barely registers.

Table 1. Sentiment Label Distribution (n=378)

Sentiment Class	n	%
Negative toward Government	156	41.3%
Positive toward Government	78	20.6%
Neutral	70	18.5%
Positive toward Activist	67	17.7%
Negative toward Activist	7	1.9%
Total	378	100%

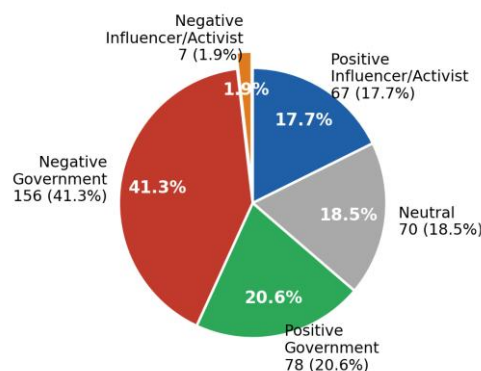


Figure 2. Sentiment Label Distribution (n=378)

The dominance of Negative toward Government (41.3%, n=156)—nearly double its positive counterpart—indicates a predominantly critical orientation in the public discourse. At the other end of the spectrum, only 7 tweets (1.9%) expressed negative sentiment toward eco-activists, small enough to suggest a consistent pattern of positive public regard rather than random variation. The resulting 22:1 majority-to-minority imbalance ratio becomes a central challenge for model training, discussed below.

For context, a naïve classifier that always predicts the most common class would achieve roughly 41.3% accuracy. KNN's 59.79% accuracy represents a genuine improvement—about 3.0 times better than random guessing under a five-class uniform baseline (20%).

KNN vs. SVM Comparison

Full per-class results are presented in Tables 2 and 3; Figures 3 and 4 complement these with visual comparisons.

Table 2. KNN Classification Report (5-Fold CV)

Class	Prec.	Recall	F1	Sup.
Neg. Activist	0.000	0.000	0.000	7
Neg. Government	0.626	0.814	0.708	156
Neutral	0.355	0.157	0.218	70
Pos. Activist	0.574	0.522	0.547	67
Pos. Government	0.639	0.680	0.658	78
Accuracy	—	—	0.598	378
Weighted avg	0.557	0.598	0.565	378

Table 3. SVM Classification Report (5-Fold CV)

Class	Prec.	Recall	F1	Sup.
Neg. Activist	0.000	0.000	0.000	7
Neg. Government	0.460	0.990	0.630	156
Neutral	0.640	0.100	0.170	70
Pos. Activist	0.680	0.190	0.300	67
Pos. Government	0.880	0.180	0.300	78
Accuracy	—	—	0.497	378
Weighted avg	0.610	0.497	0.410	378

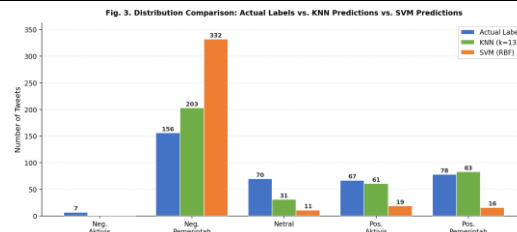


Figure 3. Distribution Comparison: Actual Label vs. KNN vs. SVM Prediction

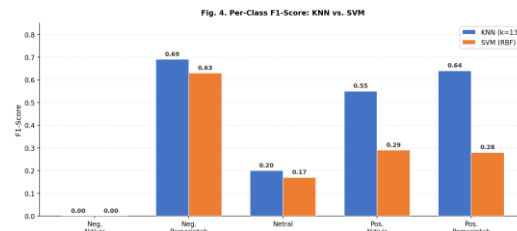


Figure 4. F1-Score per Class: KNN vs. SVM

KNN edges out SVM across the board: 59.79% accuracy and weighted F1 of 0.5651 versus 49.73% and 0.41. A more informative contrast lies in how the two models fail. SVM effectively abandons multi-class discrimination, directing 99% of its recall toward the majority class while neglecting the others—Neutral recall drops to 10%, Positive toward Government to 18%, Positive toward Activist to 19%. Fig. 3 makes this concrete: SVM predicted 332 tweets as Negative toward Government, KNN predicted 203, but the true count was only 156.

KNN distributes its attention more evenly, with recall of 81.4% for Negative toward Government, 68.0% for Positive toward Government, and 52.2% for Positive toward Activist—consistent with findings on algorithm behavior under class imbalance (Sembiring & Dewa, 2025). Neither model detects Negative toward Activist (F1=0.00 for both), unsurprising given only 7 examples exist. This implies that model choice should be guided by data characteristics rather than by prior reputation in the literature, and that a linear kernel should be considered for sparse TF-IDF representations in future work (Liu et al., 2025).

Interpretation: Institutional Skepticism and Parasocial Trust

The sentiment distribution offers one empirical entry point into two constructs mostly discussed qualitatively in the Indonesian context. This study measures sentiment polarity directed at an entity, not trust or skepticism as psychological constructs in their own right; the interpretation below is therefore consistent with, rather than confirmatory of, these frameworks. The dominance of negative government sentiment (41.3%) is suggestive of Institutional Skepticism at scale, echoing evidence that skepticism deepens when citizens perceive inadequate government engagement (Campbell, 2023) and that Indonesia's 2024 legislative campaign showed near-absent environmental commitments (Hsiao & Kuipers, 2025).

Sentiment toward eco-activists shows a markedly different pattern: only 1.9% negative versus 17.7% positive—a ratio of nearly 10:1—consistent with Parasocial Trust, though sentiment polarity alone cannot confirm the underlying psychological mechanism. Followers who engage regularly with eco-activists like Pandawara Group may develop affective bonds that formal institutions are structurally less well-placed to match (Wasike, 2025; Karnjanapoomi, 2023), a bond that congruence and self-disclosure dynamics help make durable (Koay et al., 2023a, 2023b).

Taken together, the results are consistent with civic actors being viewed as credible alongside, rather than in place of, government institutions. This should be treated as a hypothesis for future research rather than a policy conclusion: the dataset's small size ($n=378$), single-platform scope, and the indirect nature of inferring trust from sentiment polarity mean that practical implications would need testing with purpose-built instruments before being acted upon.

Limitations of the Study

This study has several limitations. First, the dataset is small (378 tweets) and confined to Twitter/X, limiting generalizability. Second, the rarest class (Negative toward Activist, $n=7$) could not be reliably learned by either classifier—a limitation of the dataset rather than the models. Third, inter-rater agreement was low (Fleiss' $\kappa = 0.13$), concentrated at the boundary between Neutral and the directional classes; future iterations should sharpen the annotation guide and re-assess agreement. Fourth, SVM used an RBF rather than a linear kernel, generally recommended for sparse TF-IDF features. Fifth, the Institutional Skepticism and Parasocial Trust framings are interpretive lenses based on sentiment polarity alone, not direct measures of trust, and should not be read as confirmed findings.

CONCLUSION AND RECOMMENDATIONS

This study mapped public sentiment patterns in Indonesian environmental discourse on Twitter/X. Negative sentiment toward the government overwhelmingly dominates the dataset (41.3%), consistent with Institutional Skepticism as an observable feature of the discourse—though, since sentiment polarity is not a direct measure of trust, this should be read as suggestive rather than confirmatory. Conversely, negative sentiment toward eco-activists stands at just 1.9%, consistent with the Parasocial Trust mechanism as one plausible explanation for the favorable sentiment attracted by figures like Pandawara Group, an interpretation that likewise awaits direct empirical testing.

On the modeling side, KNN outperforms SVM under Stratified 5-Fold Cross-Validation (59.79% vs. 49.73% accuracy; weighted F1: 0.5651 vs. 0.41). SVM's failure stems from its collapse into majority-class prediction under extreme imbalance, which inflates accuracy while destroying multi-class utility; KNN's more proportional recall makes it the more trustworthy model in this setting.

Three directions for further inquiry follow. First, the pronounced negative sentiment toward government is worth investigating further as a potential indicator that environmental communication strategy may benefit from more visible, verifiable action rather than declarative commitments alone. Second, the comparatively favorable sentiment toward eco-activists raises the question of whether they could serve a complementary role in environmental communication. Third, the five-class dual-entity labeling scheme introduced here opens a methodological pathway for comparative sentiment research in other domains where public perception of institutions and civil actors diverges. Future work should also address the study's limitations through class-imbalance techniques (e.g., SMOTE or class weighting), a linear SVM kernel, multi-platform data collection, multi-annotator labeling with reported agreement statistics, validated survey instruments for trust and skepticism, and transformer-based models such as IndoBERT.

ACKNOWLEDGEMENTS

[Optional — add acknowledgements for advisors, funding, or supporting institutions here, or delete this section if not needed.]

REFERENCES

- Adharani, S. K., Kacung, S., & Vitianingsih, A. V. (2025). Sentiment analysis on Indonesian national football team naturalization using KNN and SVM. *Edumatic*, 9(1), 189–197.
- Buvár, Á., Zsila, Á., & Orosz, G. (2023). Non-green influencers promoting sustainable consumption: Dynamic norms enhance the credibility. *Frontiers in Psychology*, 14, 1112762.
- Campbell, J. W. (2023). Public participation and trust in government: Results from a vignette experiment. *Journal of Policy Studies*, 38(2), 23–31.
- Giambastiani, G., Romito, S., & Vurro, C. (2025). When does collaborating with green influencers backfire? *Business Strategy and the Environment*, 34, 3385–3399.
- Harliyono, M., & Priskila. (2025). Social media's impact on sustainability engagement: Brands, NGOs, and influencers. *Proceedings of the Jogjakarta Communication Conference*, 3(1), 51–54.
- Hsiao, A., & Kuipers, N. (2025). Climate crisis and policy inaction in Indonesia. *American Journal of Political Science*.
- Karnjanapoomi, N. (2023). Parasocial relationships with social media influencers: Scale development and validation. *Journal of Marketing Communications*, 29(4), 412–430.
- Koay, K. Y., Cheah, C. W., & Yap, J. Y. (2023a). Self-influencer congruence, parasocial relationships, credibility, and purchase intentions. *Journal of Relationship Marketing*, 23(1), 1–20.
- Koay, K. Y., et al. (2023b). How and when social media influencers' intimate self-disclosure fosters purchase intentions. *Marketing Intelligence & Planning*, 41(6), 790–809.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.

- Liu, J., et al. (2025). Text vectorization in sentiment analysis: A comparative study of TF-IDF and Word2Vec. *ITM Web of Conferences*, 70, 03001.
- Mahmud, R. (2024). Determinants of citizen trust in public institutions in Bangladesh. *Journal of Asian and African Studies*. SAGE.
- Nazir, U., & Wani, S. A. (2024). Role of social media influencer toward environmental involvement and green buying behavior. *Business Strategy and Development*. Wiley.
- Parr, C. S. (2025). Does social media undermine trust? Institutional trust in civil society and governance institutions. *Journal of Public Policy*, 45(4), 737–760.
- Prayogo, G., & Adiwijaya, K. (2025). The influence of green influencer towards purchase intention: The extended theory of planned behavior approach. *Ilomata International Journal of Management*, 6(4), 1391–1410.
- Rohmatun, L., & Baita, A. (2025). Machine learning-based sentiment analysis on Twitter (X): A case study of the "Kabur Aja Dulu" issue using SVM. *Journal of Applied Informatics and Computing*, 9(4), 1972–1983.
- Santos, R., Coelho, P., & Marques, J. (2024). A systematic literature review on greenwashing and its relationship to stakeholders. *Management Review Quarterly*, 74, 1397–1421.
- Sembiring, A. R., & Dewa, C. K. (2025). Sentiment analysis on Indonesian tweets about the 2024 election. *Sinkron*, 9(1), 413–422.
- Sholihah, D. D., Aminah, S., & Yulistira, A. D. (2025). Building trust or masking reality? The effects of brand image, greenwashing, and influencer marketing. *Journal of Business and Economics Research (JBE)*, 6(2), 536–546.
- Uliniansyah, M. T., et al. (2024). Twitter dataset on public sentiments towards biodiversity policy in Indonesia. *Data in Brief*, 52, 109890.
- Wang, Y., Pauwels, K., & Chan, T. Y. (2023). The impact of social media influencers on sustainable consumer behavior. *Journal of Consumer Research*, 50(2), 256–273.
- Wasike, B. (2025). Me, myself, and the influencer: Examining how parasocial interaction and self-congruence with social media influencers affects news media trust. *New Media & Society*.
- Yıldırım, S., et al. (2024). The impact of influencers on sustainable consumption: A systematic literature review. *Sustainable Production and Consumption*, 52, 1–11.
- Zamsuri, A., Defit, S., & Nurcahyo, G. W. (2024). Development and comparison of multiple emotion classification models in Indonesia text. *Journal of Advanced Information Technology*, 15(4), 519–531.
- Zioło, M., Bąk, M., & Spoz, A. (2024). Literature review of greenwashing research: State of the art. *Corporate Social Responsibility and Environmental Management*, 31(6), 5343–5356.
- Zuliansyah, Z., & Jatmika, S. (2025). From screen to ground: Pandawara Group's strategy in reserving Indonesia's environment. *Jurnal Studi Komunikasi*, 9(1).