Nusantara Journal of Artificial Intelligence and Information Systems
Vol. 2, No. 1, June 2026, pp 39-46 journal homepage: https://journal.unusia.ac.id/nuai
Research Article ISSN: 3090-8302 DOI: 10.47776/nuai.v2i1.2006

Sentiment Analysis of Emotional Intensity as a Continuous Driver of Engagement and Algorithmic Visibility

Zia Ul Rehman Zafar[1*], Dedi Gunawan[2], Muhammad Saif[3]
[1,2]Department of Informatics Engineering, Universitas Muhammadiyah Surakarta Indonesia
[3]Department of Business Management, Universitas Muhammadiyah Surakarta, Indonesia

[1]l200249293@student.ums.ac.id and [2]dg163@ums.ac.id

INFORMATION

Article history:

Received March 01, 2026
Revised April 23, 2026
Accepted April 24, 2026

Keywords:

Sentiment analysis
Emotional intensity
Algorithmic visibility
Social media engagement
Digital political discourse

ABSTRACT

This study investigates how emotional intensity, rather than sentiment direction, shapes engagement and algorithmic visibility in digital political discourse. Using sentiment analysis, a dataset of about 15,000 posts from Twitter (X) and YouTube was collected over a 30-day period and scored with a hybrid TextBlob, VADER, and BERT pipeline. Emotional strength (the absolute sentiment value) correlated moderately with engagement ($r = 0.58$, $p < 0.05$), whereas the directional sentiment score did not ($r \approx 0.05$). Emotionally intense posts attracted about 2.4 times more engagement than neutral posts, and positive posts were the most frequent (41%) while neutral posts drew the lowest mean engagement. These results indicate that engagement-based ranking amplifies emotional magnitude over neutral or analytical content, which can narrow the diversity of visible expression. The findings give platform designers and policymakers a reproducible basis for assessing how affective dynamics shape visibility in algorithmically mediated public discourse.

To cite this article:

Zafar, Z. U. R., Gunawan, D., & Saif, M. (2026). Sentiment Analysis of Emotional Intensity as a Continuous Driver of Engagement and Algorithmic Visibility. *Nusantara Journal of Artificial Intelligence and Information Systems*, 2(1), 39-46.

****Corresponding Author:***

Zia Ul Rehman Zafar
Department of Informatics Engineering, Universitas Muhammadiyah Surakarta Indonesia
Taman Amir Hamzah St., Pegangsaan, Menteng, Central Jakarta, DKI Jakarta, Indonesia (10320)
Email: l200249293@student.ums.ac.id

1. INTRODUCTION

In the digital era, social media platforms have become the dominant arenas for public discourse, transforming how individuals share opinions, react to events, and participate in political life. Platforms such as Twitter (X), YouTube, Facebook, and TikTok now mediate a large share of everyday political conversation, from local civic debates to national elections [1]. Billions of posts, comments, and reactions are produced daily, and the visibility of any given message is determined less by its accuracy than by how strongly it provokes a response. Recommendation and ranking systems decide which of these messages reach a wide audience and which remain unseen, so the platform itself acts as an active filter rather than a neutral conduit [2]. This raises a basic question for digital communication: what kind of content does the engagement-based machinery of these platforms reward, and what does that mean for the diversity of voices that remain visible [3]. The answer matters not only for individual users but for the health of the wider public sphere, where unequal visibility can quietly reshape which ideas are heard.

Sentiment analysis, one of the most widely used methods in natural language processing, can quantify the emotional tone of large volumes of user text and reveal patterns that human reading alone would miss [4],[5]. The approach assigns a polarity value to a piece of text, indicating whether it leans positive, negative, or neutral, and can be applied to millions of posts at a speed no manual coder could match. Lexicon-based tools such as TextBlob [6] and VADER score text against curated word lists [7], while transformer models such as BERT capture context and handle multilingual input more robustly [8]. Recent surveys show that combining lexicon-based and deep learning models can classify sentiment and detect hate speech with practical accuracy on social media, offering a way to monitor the emotional climate of online discourse at scale [4],[9]. This makes sentiment analysis well suited to studying how emotion is distributed across large bodies of political text. Yet most of this work treats sentiment as a descriptive label rather than as a force that interacts with how platforms rank and amplify content, leaving its connection to algorithmic visibility largely unexamined [10],[11].

A growing body of evidence suggests that emotion, not neutral information, drives attention online [12]. Emotionally charged messages spread faster and further than neutral ones [13],[14], and recent audits of engagement-based ranking show that such systems amplify emotionally intense and divisive political content even when users do not prefer it [15],[3]. Because these systems are optimized to maximize interaction, content that provokes a strong reaction is structurally favored over content that informs but does not agitate [16]. This amplification has consequences for freedom of expression: when visibility is tied to emotional reaction, calm or analytical contributions are pushed to the margins, and the public sphere narrows toward its most reactive registers [17],[18]. Over time, such a dynamic can distort the perceived balance of opinion, making the loudest and most emotional positions appear more representative than they are. The result is a tension between the commercial logic of engagement and the democratic ideal of an open, balanced exchange of views.

Despite this, the relationship between the emotional properties of a message and its algorithmic visibility has not been measured directly. Prior studies tend to treat sentiment as a single linguistic artifact and rarely separate the direction of sentiment (positive versus negative) from its intensity (how strong the emotion is, regardless of direction)[19]. This conflation matters, because the two may relate to engagement in very different ways: a strongly positive post and a strongly negative post share little in direction but may behave similarly in how much attention they attract [20]. If intensity rather than direction is what platforms reward, then studies that measure only polarity will miss the mechanism that actually shapes visibility. The lack of empirical work connecting sentiment analysis to the normative question of expressive equity represents a clear research gap. Closing it requires both a measure that isolates emotional intensity and an analysis that links that measure to engagement on real platform data.

This study addresses that gap by analyzing real emotional expression and engagement across two platforms and by operationalizing emotional intensity as a measurable, continuous variable. Drawing on about 15,000 posts collected from Twitter (X) and YouTube during a 30-day window of political discussion, it scores each post for sentiment and then derives a separate measure of emotional strength from that score. Specifically, it asks whether emotional intensity predicts engagement and algorithmic visibility more strongly than sentiment direction does. The contribution is threefold: first, an operational measure of emotional intensity ($|S|$) derived directly from continuous sentiment scores and validated against human annotation; second, an empirical separation of emotional strength from sentiment direction in relation to engagement; and third, a discussion of what the observed frequency-visibility imbalance implies for fair and pluralistic digital expression. The remainder of the paper describes the data and methods, reports the sentiment distribution and correlation results, and discusses their implications for how engagement-driven platforms shape the visibility of public expression.

2. METHOD

This study uses a quantitative, cross-platform design combining sentiment analysis with correlation and sensitivity testing. Data were drawn from public political discourse on Twitter (X) and YouTube over a continuous 30-day period. The analysis pipeline has five stages: data collection, preprocessing, sentiment scoring, engagement modeling, and statistical testing. All processing was performed in Python 3.13 with a fixed random seed to support reproducibility.

2.1. Data Collection

Twitter data were gathered through the Apify Tweet Flash API, producing structured JSON output containing tweet ID, timestamp, text, and counts of likes, replies, and retweets. A parallel set of YouTube comments was retrieved with the same thematic keywords to allow cross-platform comparison. A total of 15,000 textual entries were collected (9,200 tweets and 5,800 YouTube comments). The study focused on political conversations surrounding electoral events and governance debates, since such topics generate a wide emotional spectrum and high engagement. A stratified subset (n = 2,000) was then constructed for detailed engagement analysis to ensure balanced representation across sentiment categories. After removing entries with missing engagement values during preprocessing, 1,990 posts from this subset were retained for the sentiment-engagement analysis reported in Table 2 (820 positive, 670 neutral, 500 negative).

2.2. Sentiment Scoring and Validation

This study uses lexicon-based and pre-trained sentiment models rather than supervised classifiers trained on the collected data. Consequently, no train-test split or model-training stage is required: TextBlob and VADER apply fixed lexicons [21], and BERT-base-multilingual is applied with frozen pre-trained weights [22],[6]. Sentiment polarity was computed using TextBlob, which assigns scores from -1.0 (strongly negative) to +1.0 (strongly positive). VADER was applied with its default lexicon and compound scoring, and TextBlob polarity scores were computed with its built-in analyzer without parameter modification [5],[23]. The BERT-base-multilingual-cased checkpoint was loaded through the Hugging Face transformers library [24], with input texts truncated to 512 tokens using the native WordPiece tokenizer; its output was mapped to the same three classes using the thresholds positive ($S > 0.05$), neutral ($-0.05 \leq S \leq 0.05$), and negative ($S < -0.05$) so that all three engines produce comparable labels.

$$S = \frac{1}{N}\sum_{i=1}^{N} s_i \quad (1)$$

For each text, a single sentiment score S is obtained by averaging the polarity of its sentiment-bearing tokens, as defined in Equation (1), where s_i is the polarity score of token i and N is the number of sentiment-bearing tokens in the text, so S is the mean polarity across all such tokens. The resulting score ranges from -1.0 (strongly negative) to +1.0 (strongly positive), with values near zero indicating neutral or mixed content. From this score the study derives two separate variables: the directional sentiment score S, which keeps the sign and indicates whether a post leans positive or negative, and the emotional strength |S|, the absolute value, which captures intensity regardless of direction. These two variables are analyzed separately throughout the paper because they answer different questions, one about tone and the other about magnitude. To check reliability, the automated labels were cross-validated against a stratified sample of posts annotated by two human coders ($n = 300$); inter-coder agreement reached about 92%, which reduced false positives and provided an accuracy check against model bias.

2.3. Engagement Modeling

Engagement for each post was summarized as a weighted index combining amplification signals (likes and shares) and conversational signals (comments). The weights ($\alpha = 0.4$, $\beta = 0.4$, $\gamma = 0.2$) reflect the relative role of each metric in algorithmic amplification rather than an empirically estimated parameterization; likes and shares are weighted more heavily than comments because they propagate content more directly through ranking systems. All engagement variables were normalized with min-max scaling before aggregation to allow cross-platform comparison. A sensitivity analysis under alternative weightings (equal weights, and a comment-weighted scheme) is recommended to confirm that the findings do not depend on this specific choice.

2.4. Hypotheses and Statistical Testing

The study tests one directional hypothesis. Let ρ denote the population Pearson correlation between emotional strength (|S|) and the engagement index (E). The null hypothesis (H_0) states that |S| and E are uncorrelated ($\rho = 0$), and the alternative hypothesis (H_1) states that they are positively correlated ($\rho > 0$). Because H_1 is directional, a one-tailed test is appropriate. The Pearson correlation in (2) and the associated t-statistic in (3) were used to evaluate significance, with the threshold set at $p < 0.05$. The large sample size ($n > 1{,}000$) supports the use of parametric testing under the Central Limit Theorem.

$$r_{xy} = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2 \sum_i^n (y_i - \bar{y})^2}} \quad (2)$$

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (3)$$

Before the main analysis, the reliability of the sentiment classification was checked against a human-annotated baseline, with the results reported in Table 1. The three automated engines were evaluated on the same stratified sample of 300 posts annotated by two human coders, using accuracy, precision, and recall as comparison metrics. BERT-multilingual performed best, reaching 91.2% accuracy with precision and recall above 90%, close to the 92.0% agreement observed between the human coders themselves. The two lexicon-based engines were lower but still acceptable for large-scale screening, with TextBlob at 85.6% accuracy and VADER at 82.4%. Because BERT approaches the human agreement ceiling while the lexicon models remain interpretable and fast, the study uses all three in combination, treating their convergence as evidence that a label is reliable.

Table 1. Validation of sentiment classification models against a human-annotated baseline (n = 300).

Model	Accuracy	Precision	Recall
TextBlob (lexicon)	85.6%	84.1%	83.2%
VADER (lexicon)	82.4%	81.0%	80.5%
BERT-multilingual	91.2%	90.4%	90.8%
Human agreement	92.0%	—	—

To support reproducibility, the analysis used a fixed random seed and documented preprocessing steps, and the automated sentiment labels were validated against a human-annotated baseline as reported in Table 1. The analysis code and the anonymized derived dataset are available from the corresponding author upon reasonable request, subject to platform terms of service. This allows other researchers to verify the sentiment labels, rerun the analysis pipeline, and reproduce the reported results.

3. RESULTS AND DISCUSSION

3.1. Sentiment Distribution

The overall sentiment distribution shows an uneven emotional mix in online political discourse. As shown in Table 3, positive posts formed the largest share at 41% (820 posts), followed by neutral at 33% (670 posts) and negative at 26% (500 posts). No single category formed an outright majority, which indicates that the conversation was not dominated by a single emotional tone. The sizable combined share of emotionally charged content, positive and negative together accounting for 67% of posts, reflects the polarized nature of political discussion online. This distribution sets up the central question of the analysis: whether the most frequent tone is also the one that attracts the most engagement.

The frequency of a sentiment category does not match the engagement it attracts, as Figure 1 illustrates. Although positive posts are the most frequent, neutral posts generate the lowest mean engagement (14.8), well below positive (35.2) and negative (42.7) posts. Negative sentiment attracts the highest mean engagement, suggesting that high-arousal emotional expressions are most likely to stimulate interaction. In Figure 1, this appears as a clear gap between the frequency line and the engagement bars, with negative content drawing disproportionate attention relative to its volume. This pattern points to a structural imbalance between how often content of a given tone appears and how much engagement it attracts, which is the imbalance the rest of the analysis examines.

Table 2. Sentiment distribution and mean engagement across categories (n = 1,990).

Sentiment	Posts (n)	Share (%)	Mean engagement
Positive	820	41%	35.2
Neutral	670	33%	14.8
Negative	500	26%	42.7
Total	1,990	100%	—

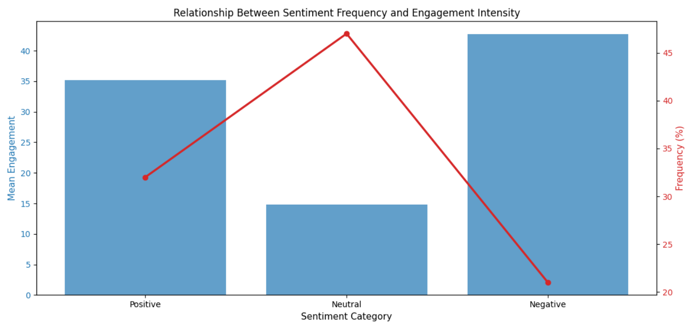


Figure 1. Relationship between sentiment frequency and mean engagement across sentiment categories.

3.2. Correlation Analysis

It is important to separate the two sentiment variables before interpreting any relationship with engagement. The directional sentiment score S correlates only weakly with engagement (r ≈ 0.05), as shown in the correlation matrix in Table 4 and visualized in the heat map in Figure 2. The emotional strength |S|, by contrast, correlates moderately and significantly with engagement ($r = 0.58$, $p < 0.05$), which points to intensity rather than direction as the relevant signal. In other words, whether a post is positive or negative has little direct association with engagement, whereas the magnitude of emotion is the stronger predictor. The coefficient of determination ($R^2 \approx 0.336$) indicates that emotional intensity alone accounts for roughly a third of the variance in engagement, leaving the remainder to factors such as topic, timing, and account reach.

Table 3. Correlation matrix of engagement metrics and the directional sentiment score (S).

	Likes	**Retweets**	**Replies**	**Sentiment (S)**
Likes	1.00	0.86	0.70	0.05
Retweets	0.86	1.00	0.86	-0.02
Replies	0.70	0.86	1.00	0.09
Sentiment (S)	0.05	-0.02	0.09	1.00

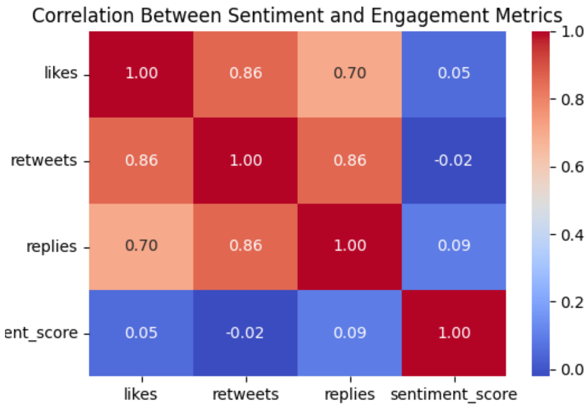


Figure 2. Correlation heat map of engagement metrics and directional sentiment score (S).

Engagement metrics confirm this relationship across the three interaction types. Posts classified as strongly positive or strongly negative received, on average, about 2.4 times more likes, replies, and shares than neutral posts. The same pattern is visible in Table 4, where likes, retweets, and replies move closely together (r between 0.70 and 0.86) while sentiment direction barely registers. This result supports rejecting the null hypothesis (H_0), confirming that emotional intensity is positively associated with engagement. Taken together, the matrix and the heat map in Figure 2 show that strong correlations appear among likes, retweets, and replies, while S shows only weak associations with all three, confirming that sentiment direction alone is a poor predictor of engagement. Audiences respond to how strongly a post is charged, not to whether its tone is favorable or critical.

3.3. Sensitivity Analysis

The engagement index combines likes, shares, and comments using fixed weights ($\alpha = 0.4$, $\beta = 0.4$, $\gamma = 0.2$), chosen to reflect the relative role of each signal in algorithmic amplification rather than estimated from the data. Because this choice is not empirically derived, a sensitivity analysis that perturbs these weights and recomputes the correlation is needed to confirm that the central finding is not an artifact of the weighting. Such an analysis would compare the baseline against equal-weight and comment-weighted schemes and check whether the direction and significance of the emotional-strength–engagement relationship hold. We identify this as a limitation of the present study and a priority for future work. Reporting the correlation under several weighting schemes would make the robustness of the result transparent and verifiable.

3.4. Discussion

The findings reveal a clear pattern in algorithmically mediated discourse: the intensity of emotion, rather than its direction, governs public attention online. This is most visible in the contrast between the two sentiment variables, where emotional strength correlated moderately with engagement ($r = 0.58$) while the directional sentiment score did not ($r \approx 0.05$). Posts that express strong feeling, whether positive or negative, were systematically more visible, receiving about 2.4 times more engagement than neutral posts. Neutral and analytical contributions drew the least amplification, with a mean engagement of 14.8 against 35.2 for positive and 42.7 for negative posts. This is consistent with prior work showing that emotionally charged and morally laden content diffuses more widely [12],[13],[14] and that engagement-based ranking amplifies divisive material [15], [3].

A second observation is the mismatch between how often a tone appears and how much attention it receives. Although positive posts were the most frequent category at 41% of the sample, negative posts, the least common at 26%, attracted the highest mean engagement. This frequency-visibility gap means that the emotional profile of what circulates widely does not reflect the emotional profile of what is actually posted. The directional sentiment of a message, then, is a weak guide to its reach, since two posts of opposite tone but similar intensity can perform alike. The mechanism that shapes visibility is the magnitude of emotion, not whether the sentiment is favorable or critical.

These dynamics carry implications for freedom of expression. When visibility is conditioned on emotional reaction, the platform environment rewards a narrow band of expressive styles and can crowd out measured or minority voices, contributing to a thinner and more polarized public sphere [16],[17],[18]. The concern is not that emotion is present in political talk, which is normal and legitimate, but that its magnitude becomes a precondition for being heard. This raises questions of expressive equity that engagement counts and similar content metrics do not capture. If the calmest and most analytical contributions are structurally the least visible, the cost is borne not only by individual users but by the quality of collective deliberation. Addressing this would require ranking signals that account for contribution quality alongside reaction, rather than reaction alone.

The study also has limitations that bound these conclusions. The dataset is drawn from two platforms over a 30-day window, which limits observation of longer-term trends and platform-specific effects. Engagement metrics capture interaction quantity but not the qualitative depth of discourse, and lexicon-based scoring can struggle with sarcasm, dialect, and contextually appropriate but emotionally charged language [25],[26]. The engagement index also relies on fixed weights that were not estimated from the data, so the precise magnitude of the correlation should be read as indicative rather than exact. The sentiment labels are validated through cross-method consistency and human agreement rather than against measured real-world outcomes such as long-term influence on opinion. These limitations point to clear next steps: multilingual and cross-cultural data, longer observation windows, a sensitivity check on the engagement weighting, and validation against downstream outcomes.

4. CONCLUSION

This study examined how natural language processing can quantify the relationship between emotional expression and engagement in digital political discourse, and it operationalized emotional intensity as a continuous, measurable variable. Across about 15,000 posts from Twitter and YouTube, emotional strength correlated moderately with engagement ($r = 0.58$, $p < 0.05$) while directional sentiment did not, and emotionally intense content drew roughly 2.4 times more engagement than neutral content. Positive posts were most frequent, yet neutral posts attracted the least engagement, exposing a frequency-visibility imbalance. Together these results indicate that engagement-based ranking amplifies emotional magnitude over neutral or analytical content, which can narrow the diversity of visible expression. The main contribution is a transparent way to separate emotional intensity from sentiment direction and to link it empirically to algorithmic visibility. Future work should extend the analysis to multilingual and cross-cultural datasets, incorporate economic and contextual factors, validate rankings against measured outcomes, and explore hybrid sentiment and machine-learning approaches for assessing expressive equity in algorithmically mediated communication.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this paper.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] S. Gearhart, B. Zhang, and O. Adegbola, "Tweeting, talking, or doing politics? Testing the influence of communication on democratic engagement," *Telemat. Informatics Reports*, vol. 16, p. 100167, 2024, doi: https://doi.org/10.1016/j.teler.2024.100167.

[2] T. Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. 2018. doi: 10.12987/9780300235029.

[3] P. Bouchaud, "Algorithmic Amplification of Politics and Engagement Maximization on Social Media," 2024, pp. 131–142. doi: 10.1007/978-3-031-53503-1_11.

[4] M. Subramanian, S. Veerappampalayam Easwaramoorthy, G. Deepalakshmi, J. Cho, and G. Manikandan, “A survey on hate speech detection and sentiment analysis using machine learning and deep learning models,” *Alexandria Eng. J.*, vol. 80, pp. 110–121, Oct. 2023, doi: 10.1016/j.aej.2023.08.038.

[5] B. Pang and L. Lee, “Opinion Mining and Sentiment Analysis,” *Found. Trends Inf. Retr.*, vol. 2, pp. 1–135, Jul. 2008, doi: 10.1561/1500000011.

[6] W. Aljedaani *et al.*, “Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry,” *Knowledge-Based Syst.*, vol. 255, p. 109780, 2022, doi: https://doi.org/10.1016/j.knosys.2022.109780.

[7] K. Barik and S. Misra, “Analysis of customer reviews with an improved VADER lexicon classifier,” *J. Big Data*, vol. 11, no. 1, p. 10, 2024, doi: 10.1186/s40537-023-00861-x.

[8] T. Islam *et al.*, “Lexicon and Deep Learning-Based Approaches in Sentiment Analysis on Short Texts,” *J. Comput. Commun.*, vol. 12, pp. 11–34, Jan. 2024, doi: 10.4236/jcc.2024.121002.

[9] F. Alkomah and X. Ma, “A Literature Review of Textual Hate Speech Detection Methods and Datasets,” *Information*, vol. 13, p. 273, May 2022, doi: 10.3390/info13060273.

[10] T. Garg, S. Masud, T. Suresh, and T. Chakraborty, “Handling Bias in Toxic Speech Detection: A Survey,” *ACM Comput. Surv.*, vol. 55, Jan. 2023, doi: 10.1145/3580494.

[11] A. Albladi *et al.*, “Hate Speech Detection Using Large Language Models: A Comprehensive Review,” *IEEE Access*, vol. PP, p. 1, Jan. 2025, doi: 10.1109/ACCESS.2025.3532397.

[12] J. Berger and K. Milkman, “Emotion and Virality: What Makes Online Content Go Viral?,” *GfK Mark. Intell. Rev.*, vol. 5, pp. 18–23, May 2013, doi: 10.2478/gfkmir-2014-0022.

[13] S. Stieglitz and L. Dang-Xuan, “Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior,” *J. Manag. Inf. Syst.*, vol. 29, no. 4, pp. 217–248, Apr. 2013, doi: 10.2753/MIS0742-1222290408.

[14] W. J. Brady, J. A. Wills, J. T. Jost, J. A. Tucker, and J. J. Van Bavel, “Emotion shapes the diffusion of moralized content in social networks,” *Proc. Natl. Acad. Sci.*, vol. 114, no. 28, pp. 7313–7318, Jul. 2017, doi: 10.1073/pnas.1618923114.

[15] S. Milli, M. Carroll, Y. Wang, S. Pandey, S. Zhao, and A. Dragan, “Engagement, user satisfaction, and the amplification of divisive content on social media,” *PNAS Nexus*, vol. 4, Mar. 2025, doi: 10.1093/pnasnexus/pgaf062.

[16] S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books, 2019. [Online]. Available: https://books.google.co.id/books?id=bYL9xAEACAAJ

[17] M. Haim, A. Graefe, and H.-B. Brosius, “Burst of the Filter Bubble?: Effects of personalization on the diversity of Google News,” *Digit. Journal.*, vol. 6, pp. 330–343, Mar. 2018, doi: 10.1080/21670811.2017.1338145.

[18] J. Habermas, *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. Cambridge: The MIT Press, 1991.

[19] K. L. Keung, C. K. M. Lee, and K.-T. Luk, “Determinants of Behavioral Intention in Augmented Reality Filter Adoption: An Integrated TAM and Satisfaction–Loyalty Model Approach,” *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 20, no. 3. p. 186, 2025. doi: 10.3390/jtaer20030186.

[20] P. Nandwani and R. Verma, “A review on sentiment analysis and emotion detection from text,” *Soc. Netw. Anal. Min.*, vol. 11, no. 1, p. 81, 2021, doi: 10.1007/s13278-021-00776-6.

[21] C. J. Hutto and E. Gilbert, *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text*. 2015.

[22] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *NAACL-HLT*, Oct. 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.

[23] J. Zhao, K. Liu, and L. Xu, “Sentiment Analysis: Mining Opinions, Sentiments, and Emotions,” *Comput. Linguist.*, vol. 42, pp. 595–598, Sep. 2016, doi: 10.1162/COLI_r_00259.

[24] A. Aljabar, I. Ali, and B. M. Karomah, “Sentiment Analysis Using Transformer Method,” *J. Informatics Inf. Syst. Softw. Eng. Appl.*, vol. 6, no. 2, pp. 90–97, 2024, doi: 10.20895/insista.v6i2.1383.

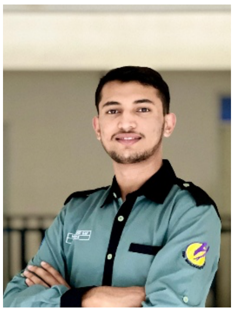
[25] Z. Waseem and D. Hovy, *Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter*. 2016. doi: 10.18653/v1/N16-2013.

[26] S. Mohammad and P. Turney, “Crowdsourcing a Word-Emotion Association Lexicon,” *Comput. Intell.*, vol. 29, Aug. 2013, doi: 10.1111/j.1467-8640.2012.00460.x.

BIOGRAPHIES OF AUTHORS

Zia Ul Rehman Zafar is an undergraduate student in the Department of Informatics Engineering at Universitas Muhammadiyah Surakarta, Indonesia, under the International Priority Scholarship program. His research interests include natural language processing, artificial intelligence ethics, machine learning, and digital governance. He has participated in international academic conferences and actively engages in interdisciplinary research integrating technology with social responsibility. His current work explores ethical AI frameworks for digital discourse and content moderation. He can be contacted at email:

Dedi Gunawan is a senior academic and researcher at Universitas Muhammadiyah Surakarta, Indonesia, where he serves as Head of the Informatics Engineering Department and Program Coordinator for both the Informatics Undergraduate and Master's programs. He holds a doctoral degree (Ph.D.) and has extensive experience in computer science education, academic leadership, and research supervision. Dr. Gunawan has represented UMS in global research initiatives such as the JST Nexus Green AI collaboration with NAIST (Nara Institute of Science and Technology), Japan, where he also served as a keynote speaker alongside colleagues from other international institutions. He can be contacted at email: dg163@ums.ac.id



Muhammad Saif is an undergraduate student in the Department of Business Management at Universitas Muhammadiyah Surakarta, Indonesia. His academic interests include digital business strategy, technology governance, and the socio-economic implications of artificial intelligence. He is particularly interested in exploring how emerging technologies influence digital communication, entrepreneurship, and sustainable development.