

Analysis of the Influence of Social Media on Generation X in Choosing a North Sumatra Governor Candidate

Alberto^{1*}, Pieter Octaviandy², Christina NM Tobing³

^{1,2,3}Information Systems, STMIK Time, Meda, Indonesia

albertolamm4118@gmail.com^{1*}, pieter.lecture@gmail.com², christinamtobing@gmail.com³

Abstract

The dynamics of digital politics, particularly in regional head elections, are heavily influenced by the flow of information on online and social media. This study examines how online content shapes the political perspectives and preferences of Generation X within the context of the North Sumatra gubernatorial election (Pilgub Sumatera Utara), a key voter segment. Using a mixed-methods approach, this research analyzed 414 text entries from online news articles and social media comments, as well as questionnaire data from 117 Gen X respondents. Sentiment analysis revealed that online sentiment was predominantly Neutral (31.62%), followed by Negative (29.91%) and Positive (20.51%). Although Generation X actively uses social media for information (median 4.0), the direct influence of public comments was rated low (median 2.0). "Online articles," however, were found to be the most influential source (41.88%). While their trust in social media information is moderate (median 3.0), Gen X's candidate support preference is clear (61.54%), suggesting that their choices are formed through a critical evaluation of information sources rather than passive consumption. The methodology proved effective in providing comprehensive insights into digital public opinion and the perspective of Generation X.

Keywords: Social Media, Generation X, North Sumatra Gubernatorial Election, Sentiment Analysis, Public Opinion.

1. Introduction

The political landscape in the digital era has significantly altered regional head elections in Indonesia, with online and social media platforms now serving as primary arenas for information dissemination and campaigning. In the context of the 2024 North Sumatra gubernatorial election (Pilgub Sumatera Utara), where the Permanent Voter List (DPT) reached 10,771,496, Generation X (born 1965-1980) represents a significant voter segment. Numbering 2,723,795, they account for 25.29% of the total voters and have a high social media access rate (66.3% on Facebook).

Despite having decades of voting experience, the proliferation of political content on social media—including unverified information, opinions, or hoaxes—raises a crucial question about its potential to influence perspectives and change political choices. This phenomenon highlights the urgency for voters to be wise in processing digital information and poses a challenge for the government in educating the public.

This research aims to understand how various types of political content on social media influence the perspectives and political choices of Generation X in the North Sumatra gubernatorial election. Specifically, this study will analyze the patterns of public sentiment (positive, negative, neutral) in online news articles (from Detik.com, Kompas, Tempo, CNN) and social media comments (from Reddit, X, Facebook, Instagram). It will also examine the correlation between Generation X's trust in Pilgub information on social media and their preference for a gubernatorial candidate. By using a mixed-methods approach that combines Natural Language Processing (NLP)-based sentiment analysis with the Naive Bayes method on text data and descriptive statistical analysis on Generation X respondent questionnaire data, this research is expected to provide comprehensive insights into the dynamics of digital public opinion and the voting behavior of Generation X.

2. Page layout

2.1 Social Media and Digital Politics

Social media platforms are a crucial tool for communication and information exchange, allowing users to create, share, and consume content. The term "media" refers to communication channels, while "social" relates to shared information and awareness, emphasizing social media's role in connecting individuals virtually. In politics, these platforms have become a key arena for elections, facilitating candidate promotion and the spread of political information. However, this accessibility also poses significant challenges, particularly the rapid spread of misinformation and "hoaxes," which can heavily influence public perception and political dynamics.

2.2 Generation X as Digital-Era Voters

A "generation" is defined as a group of individuals sharing a common birth year, location, and significant historical experiences. Generation X, born between 1965 and 1980, serves as a bridge between Baby Boomers and Millennials/Gen Z. Members of this generation, who will be aged 44 to 59 in 2024, are known for their resilience and high adaptability to rapid change, having transitioned from the pre-digital to the current digital era. They grew up amidst technological revolutions like the emergence of personal computers and the internet and tend to be independent and skeptical of authority. Data indicates that about 70% of Gen X uses social media for information, including political news. They generally use these platforms for practical, informational purposes, such as getting news and discussing specific topics, and are less susceptible to the "filter bubble" common among younger generations. This unique position makes Gen X a crucial voter segment to study in the context of how social media influences political perspectives, especially in regional elections.

2.3 Text Analysis Methods and Political Data Classification

Human digital activity generates massive volumes of textual data, a key component of big data characterized by volume, velocity, variety, veracity, and value. Natural Language Processing (NLP) is a vital technique for extracting knowledge from this data, allowing text to be processed automatically, similar to how humans analyze it. In political communication research, NLP is essential for tasks like text classification, which groups documents into specific categories, such as in sentiment analysis or hate speech detection.

Human digital activity generates massive volumes of textual data, a key component of big data characterized by volume, velocity, variety, veracity, and value. Natural Language Processing (NLP) is a vital technique for extracting knowledge from this data, allowing text to be processed automatically, similar to how humans analyze it. In political communication research, NLP is essential for tasks like text classification, which groups documents into specific categories, such as in sentiment analysis or hate speech detection.

3. Result and Discussion

3.1 Research Design

This study uses a mixed-methods approach, integrating computational analysis of text data with a questionnaire survey, to examine the influence of social media on the perspectives of Generation X in the 2024 North Sumatra gubernatorial election. This design allows for a broad sentiment analysis of public opinion on online media, while also understanding the specific perceptions and support preferences of the Generation X voter segment.

3.2 Data Collection

Primary data collection was conducted through two main approaches:

a. Online and Social Media Text Data

Textual data was collected using a web scraping program developed with Python. The data was sourced from various leading online news portals (Detik.com, Kompas, Tempo, CNN) and social media platforms (X/Twitter, Facebook, Instagram, Reddit). The collected data included news articles, comments, tweets, and public posts relevant to the 2024 North Sumatra gubernatorial election, identified through the use of hashtags or keywords related to the candidates (e.g., "#PemiluSumut2024", "Bobby", "Edy"). The textual data was filtered for topic relevance and language (Indonesian only).

b. Generation X Questionnaire Data

To obtain specific perspectives from Generation X, an online questionnaire was distributed via Google Forms. The questionnaire was designed to measure the frequency of social media use for political information, the level of trust in gubernatorial election information on social media, perceptions of the gubernatorial candidates, and the influence of social media information on respondents' political choices. The influence was measured using a Likert scale (1=very little influence, 5=very high influence). The target respondents were Generation X (born 1965–1980) in North Sumatra, with at least 50 respondents who are active on social media or exposed to gubernatorial election information. A purposive sampling technique was used, and the questionnaire included an automatic filter based on year of birth to ensure the validity of the respondents' demographics.

3.3 Data Analysis

The data collected from both sources was analyzed using the following approaches:

a. Text Data Pre-processing

Text data from online and social media underwent a series of pre-processing stages to prepare it for analysis. This included the removal of special characters, links (URLs), mentions, hashtags, and emoticons; removal of duplicate data; conversion of text to lowercase; and word normalization (e.g., stemming). This process was aided by Python libraries such as NLTK or Sastrawi for handling Indonesian text.

b. Sentiment Analysis with Naive Bayes

The Naive Bayes algorithm was used as the primary method for sentiment classification of the pre-processed textual data. Sentiment was categorized as positive, negative, or neutral towards the gubernatorial candidates. A portion of the text data was manually labeled (positive, negative, neutral) to train the Naive Bayes model. The trained model was then used to classify the sentiment of all unlabeled textual data. The model's performance was evaluated using standard metrics such as accuracy, precision, recall, and F1-score to ensure the reliability of the classification. The results of the analysis were visualized using a distribution graph.

c. Descriptive Statistical Analysis of Questionnaire Data

The questionnaire responses were analyzed using descriptive statistics. This included calculating the frequency and percentage of each answer choice, especially for the Likert scale questions, to identify trends in perception and the level of influence.

d. Data Integration

The results from the textual data sentiment analysis and the descriptive statistical analysis of the questionnaire data will be integrated. This integration aims to compare the dominant sentiment patterns in the online public sphere with the trust levels, perceptions, and support

preferences of Generation X towards the gubernatorial candidates, thereby providing a comprehensive overview of the influence of social media.

4. A step before the final submission

This section presents the main findings of the study from the textual data and questionnaire analyses, followed by an integrated interpretation.

a Online and Social Media Sentiment Analysis

The Naive Bayes sentiment analysis showed variations in the model's pattern and performance. For Detik.com articles (112 Articles), the dominant sentiment was Positive (71.4%), followed by Neutral (25.0%), and Negative (3.6%).

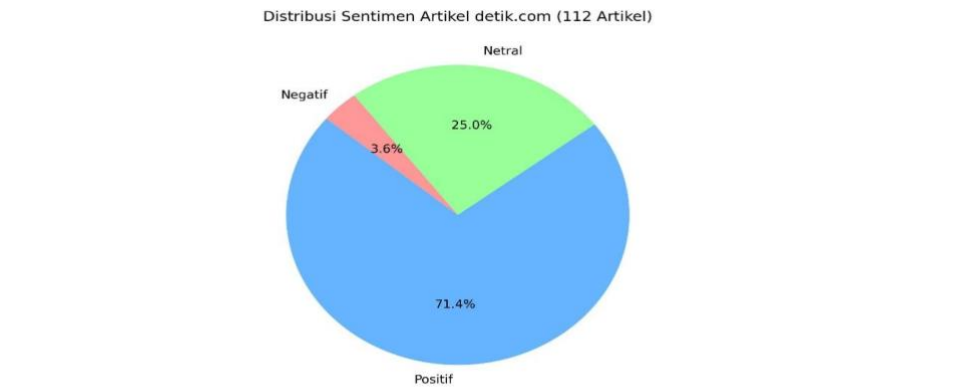


Fig. 1: Sentiment Distribution of Detik.com Articles

The model for this dataset achieved an accuracy of 0.82.



Fig. 2: Detik.com Model Classification Report

For the combined/manual articles, sentiment was also predominantly Positive (70.1%), with Neutral (16.5%), and Negative (13.4%).

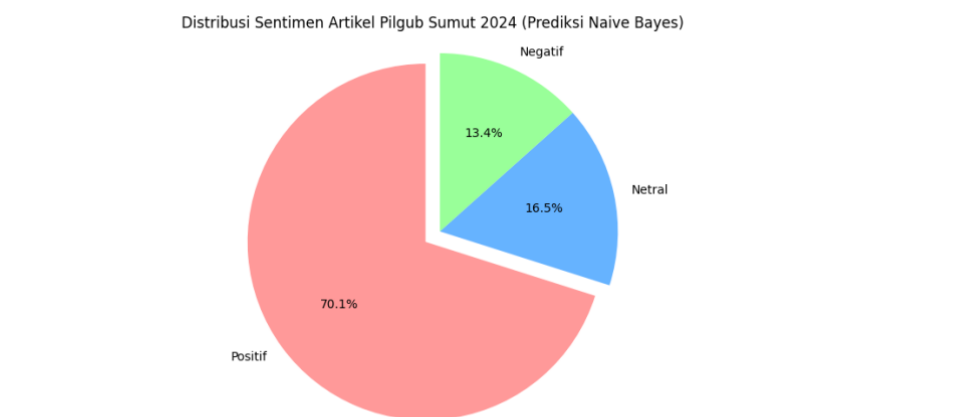


Fig. 3: Sentiment Distribution of Combined/Manual Articles

The model's accuracy for this data was slightly higher, at 0.85.

```
Laporan Klasifikasi:
      precision    recall  f1-score   support

   Negatif      0.67      1.00      0.80         2
    Netral      0.75      0.75      0.75         4
    Positif      0.92      0.86      0.89        14

 accuracy      0.85         20
macro avg      0.78      0.87      0.81         20
weighted avg   0.86      0.85      0.85         20

Hasil prediksi disimpan di: articles_pilgub_sumut_labeled_100_predicted.csv
```

Fig. 4: Combined/Manual Model Classification Report

In contrast, social media comments (205 rows) showed a more varied distribution: Positive (47.8%), Negative (29.8%), and Neutral (22.4%).

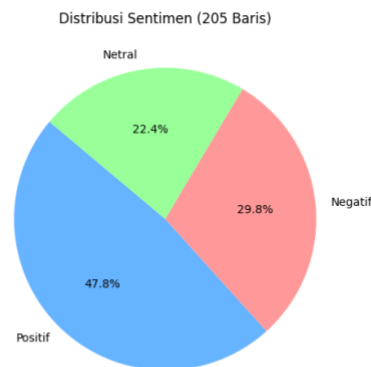


Fig. 5: Sentiment Distribution of Social Media Comments

The model's accuracy for these comments was lower (0.63), reflecting the complexity of language on social media.

```
      precision    recall  f1-score   support

   Negatif      0.67      0.33      0.44        12
    Netral      0.00      0.00      0.00         6
    Positif      0.63      0.96      0.76        23

 accuracy      0.63         41
macro avg      0.43      0.43      0.40         41
weighted avg   0.55      0.63      0.56         41

Model disimpan di: C:/Users/user/OneDrive/Desktop/naive_bayes_model_205.pkl
Vectorizer disimpan di: C:/Users/user/OneDrive/Desktop/efidf_vectorizer_205.pkl

Distribusi Sentimen pada Data yang Dilabeli (205 baris):
manual_sentiment
Positif      98
Negatif      61
Netral       46
Name: count, dtype: int64
>> |
```

Fig. 6: Social Media Comments Model Classification Report

Overall, the majority sentiment of respondents on social media (based on the questionnaire) was dominated by Neutral (31.6%) and Negative (29.9%), with Positive at 20.5% and Don't Know at 17.9%. This indicates a more diverse public perception compared to the news narrative.

b Generation X Questionnaire Analysis

The questionnaire (117 Generation X respondents) showed that they frequently use social media for gubernatorial election information (median 4.0). However, their level of trust tends to be neutral (median 3.0), and the influence of public social media comments on their choice is low (median 2.0). This implies that Generation X has a strong information filter. Bobby Nasution received majority support (61.54%), far ahead of Edy Rahmayadi (18.80%). Besides social media, online articles were the most dominant source of information (41.88%).

c Integration of Findings and Answers to Research Questions

The influence of social media posts on Generation X is complex: they actively use it, but personal filters result in a low direct influence from public comments (RQ1). The sentiment patterns vary from predominantly positive in news to more diverse and even negative on social media from the respondents' perspective (RQ2). Although the level of trust in social media information is moderate, Generation X's support preference is very clear, indicating that their decisions are influenced by a combination of trusted sources (such as online articles) and effective candidate branding, not merely by general sentiment on social media (RQ3).

d Strengths, Limitations, and Suggestions for Development

This research excels in its mixed-methodology and focus on Generation X, but is limited by the scope of data and the nature of the Naive Bayes model. Suggestions for future development include expanding data sources, using more sophisticated NLP models, conducting in-depth thematic issue analysis, more detailed respondent segmentation, longitudinal studies, and developing an interactive dashboard.

5. Conclusion

This research aims to analyze public sentiment on online and social media regarding the North Sumatra gubernatorial election, and to examine the influence of social media posts on the perspectives and support preferences of Generation X. Based on a comprehensive data analysis from web scraping and a questionnaire, several key conclusions were drawn:

1. The Influence of Social Media Posts on Generation X's Perspective

Generation X shows a high level of engagement in using social media as a source of gubernatorial election information (median 4.0 'Often'). However, they tend to be less influenced by direct public comments on social media when making their choice (median 2.0 'Slight Influence'). Although 41.88% of Generation X relies on "online articles" as the most influential source of information, the sentiment from these articles (which are also shared on social media) can indirectly shape or strengthen their perspective. This indicates that the influence of social media on Generation X's perspective is indirect and complex, working more through exposure to news content than through direct persuasion from comments.

2. Public Sentiment Patterns

The sentiment patterns identified on online and social media are dominated by Neutral sentiment (31.62%), followed significantly by Negative sentiment (29.91%), then Positive (20.51%), and the rest being "Don't Know" (17.95%). This suggests that the digital discourse surrounding the North Sumatra gubernatorial election tends to be informative and balanced, but with a significant portion of criticism. The difference in sentiment between news articles (which tend to be positive/neutral) and social media comments (which are more volatile and emotional) indicates the distinct role of each platform in shaping public opinion.

3. Correlation between Generation X's Trust Level and Support Preference

Generation X shows a moderate or neutral level of trust (median 3.0) towards general gubernatorial election information on social media, indicating a cautious and critical attitude. Nevertheless, there is a very clear support preference for Bobby Nasution (61.54%). This correlation implies that Generation X's support preference is not solely driven by a general trust level in all social media information, but rather by factors such as the credibility of specific information sources (e.g., online articles), effective candidate narratives, or offline factors that successfully penetrate their critical filter.

4. Suggestions for Future Researchers and Campaign Teams

For future researchers, it is recommended to expand the scope of social media data, improve the accuracy of sentiment models with more advanced NLP, and conduct in-depth thematic issue analysis. Longitudinal studies and more specific respondent segmentation (based on demographics/psychographics) could also provide richer insights.

For campaign teams or candidates, it is recommended to optimize digital communication strategies by developing proactive and positive narratives, focusing on the dissemination of informative and credible content (especially online articles), and building authentic interactions relevant to Generation X.

References

- [1] R. A. Prasetya, "PENGARUH POLITIK IDENTITAS MELALUI MEDIA SOSIAL TERHADAP GENERASI MILENIAL DAN PELAKSANAAN PEMILU," dalam *Conference On Communication and News Media Studies*, 2019, p. 21.
- [2] B. Arianto, "Salah Kaprah Ihtwal Buzzer: Analisis Percakapan Warganet di Media Sosial," *JIP: Jurnal Ilmiah Ilmu Pemerintahan*, vol. 5, no. 1.
- [3] N. Kadir, "Media Sosial dan Politik Partisipatif : Suatu Kajian Ruang Publik, Demokrasi Bagi Kaum Milenial dan Gen Z," *RESIPROKAL: Jurnal Riset Sosiologi Progresif Aktual*, vol. 4, no. 2, 2022.
- [4] M. Y. Cobis and U. Rusadi, "Sosial Media Sebagai Media Kampanye Partai Politik," *Da'watuna: Journal of Communication and Islamic Broadcasting*, vol. 3, no. 4, 2023.
- [5] R. S. R. A. Kartodinoto, D. Purnomo, and A. W. Utomo, "Analisis Instagram Sebagai Ruang Publik Terhadap Respon Generasi-Z Dalam Kontestasi Pilpres 2024 (Studi Pada Generasi-Z Di Kota Salatiga)," Universitas Kristen Satya Wacana.
- [6] D. O. Andriyendi, N. S., and S. F. Dewi, "Media sosial dan pengaruhnya terhadap partisipasi politik pemilih pemula pada Pilkada," *Journal of Education, Cultural and Politics*, vol. 3, no. 1, 2023.
- [7] D. L. Y. Lopulalan, *Teori-Teori Komunikasi*.
- [8] Y. Rusfiana and I. Nurdin, *Dinamika Politik Kontemporer*. Bandung: Alfabeta, 2017.
- [9] S. Anggara, *Sistem Politik Indonesia*. Bandung: Pustaka Setia, 2013.
- [10] F. Fitriyah, *Teori dan Praktik Pemilihan Umum di Indonesia*. Yogyakarta.
- [11] KISP, "PERILAKU MILLENNIAL VOTERS DI MEDIA SOSIAL DALAM MERESPON PEMILIHAN UMUM SERENTAK 2019."
- [12] Media Culture, "Generasi Yang Cerdas Secara Digital : Bagaimana Gen X Merangkul Teknologi."
- [13] R. Y. Hayuningtyas, "Aplikasi Filtering of Spam Email Menggunakan Naïve Bayes."
- [14] I. A. Rahma and L. H. Suadaa, "PENERAPAN TEXT AUGMENTATION UNTUK MENGATASI DATA YANG TIDAK SEIMBANG PADA KLASIFIKASI TEKS BERBAHASA INDONESIA."