



Mapping Public Opinion on the DPR Salary Increase Issue via YouTube Comment Sentiment Analysis using IndoBERTa

Rifki Saefullah^{1*}, Mugi Lestari²

^{1,2}*Communication in Research and Publications, Bandung, Indonesia*

**Corresponding author email: rifki.saefullah6@gmail.com*

Abstract

The issue of salary increases for members of the House of Representatives (DPR) in Indonesia has sparked widespread public debate because it is considered inconsistent with the socio-economic conditions of the community. The policy is perceived as sensitive, particularly regarding the principles of social justice, government accountability, and political legitimacy. In the digital era, social media such as YouTube has become an important space for the public to express opinions openly. This study aims to map public opinion regarding the DPR salary increase issue through sentiment analysis of YouTube comments using a Natural Language Processing (NLP) approach. The IndoBERTa model was used to classify public sentiment into positive, negative, and neutral categories, and n-gram analysis was used to capture dominant linguistic patterns. The results showed that negative sentiment dominated with 5,463 comments, far exceeding neutral (1,391 comments) and positive (812 comments). The n-gram analysis revealed that frequently appearing words and phrases related to "people," "DPR," "salary," as well as emotional expressions such as "disband the DPR" and "dancing on the people's suffering." These findings indicate that the DPR salary issue triggered a strong, often sarcastic public response and demonstrated a crisis of trust in the legislative institution. The lack of positive sentiment confirms that this policy has almost no public support.

Keywords: Public sentiment, dpr, dpr member salaries, indoberta, youtube

1. Introduction

The issue of salary increases for members of the Dewan Perwakilan Rakyat (DPR) in Indonesia has sparked widespread public debate. This discourse is considered sensitive because it concerns the use of the state budget and public trust in legislative performance. For some, DPR salary increases are considered irrelevant to the still challenging socio-economic conditions, such as high poverty rates, inequality, and budgetary needs in priority sectors like health and education.

The debate over DPR salary increases is not merely an administrative issue, but also concerns the principles of social justice and government accountability. Ideally, the salaries and allowances of representatives should align with performance, transparency, and contributions to public welfare (Marcinkowski et al., 2024). When DPR salary policies are deemed out of balance with the real conditions of society, negative perceptions emerge that can weaken political legitimacy and public trust in the legislative institution.

In today's digital era, social media has become a crucial platform for channeling public aspirations. YouTube, as one of the largest platforms, plays a significant role in the dissemination of political information and serves as a forum for public discussion (Isibor et al., 2021; Swastiningsih et al., 2024). Comment sections on videos related to political issues, including the Dewan Perwakilan Rakyat salary increase, are often filled with diverse public opinions, ranging from support and criticism to sarcasm. This phenomenon demonstrates that YouTube can be a potential data source for understanding public opinion in real time.

Sentiment analysis has emerged as an important method for processing public comment data on social media. With this approach, public opinion can be categorized as positive, negative, or neutral, allowing researchers to understand public response patterns on specific issues. Machine learning-based sentiment analysis techniques can also help present data more objectively and systematically than conventional survey methods.

Developments in Natural Language Processing (NLP) are driving the use of advanced language models in public opinion analysis. One relevant model is IndoBERTa (Indonesian BERT-based Sentiment Classifier) which was developed to process multilingual text more accurately. This model enables the analysis of YouTube comments in

Indonesian with more contextual results, as it is able to understand complex linguistic nuances (Ojadi et al., 2023; Alam et al., 2025).

The use of IndoBERTa in sentiment analysis related to the Dewan Perwakilan Rakyat salary increase issue is crucial because public comments on YouTube often contain informal language, mixed language, and emotional expressions. This model can capture semantic representation better than traditional methods, resulting in more accurate and relevant public opinion mapping results (Babu & Eswari, 2021).

This study emphasizes the importance of mapping public opinion on the Dewan Perwakilan Rakyat salary increase policy through a sentiment analysis approach. By utilizing YouTube comment data, the study can provide a realistic picture of public response, both in support and opposition. This information can be used as evaluation material for policymakers and legislative institutions to increase transparency and responsiveness to public opinion. Furthermore, this study also provides a methodological contribution to the development of social media data analysis in Indonesia. Public opinion studies have been dominated by traditional surveys, which are expensive and time-consuming. With a big data and NLP approach, public opinion analysis can be conducted quickly, massively, and sustainably, better adapting to the dynamics of ongoing digital discourse (Klingenberg et al., 2022; Stefanovic et al., 2025).

Mapping public opinion on the DPR salary increase through sentiment analysis can also enrich the literature on digital politics in Indonesia. The results of this study are expected to explain how people express political attitudes in the virtual space, while also demonstrating the relationship between public policy and social legitimacy in the digital era. Therefore, this study aims to map public opinion on the DPR salary increase issue based on YouTube comments using the IndoBERTa model. The results are expected to provide a more comprehensive understanding of public perceptions and contribute to the development of more effective, data-driven political communication strategies.

2. Research Methods

This study uses a quantitative approach based on text analysis to understand public opinion regarding the issue of Dewan Perwakilan Rakyat (DPR) salary increases through comments on the YouTube platform. Comment data was collected from ten videos relevant to the topic using the *youtube – comment – downloader* library. Each collected comment was stored along with metadata information, such as the video URL, the commenter's name, the comment's time, the text content, and the number of votes received. The total number of comments collected reached several thousand entries, which formed the basis for this research database.

The raw comment data was then processed using a series of text cleaning steps to improve the quality of the analysis. This process included converting the text to lowercase, removing URLs, numbers, punctuation, and excess spaces. The goal was to minimize noise in the n-gram and sentiment analysis, so that the results obtained were more representative of actual public opinion.

The n-gram analysis stage was conducted to identify the most frequently occurring words or phrases in the comments. Using *CountVectorizer* from scikit-learn, unigrams (single words) and bigrams (two-word combinations) were extracted from the entire cleaned comment text. This analysis provides initial insight into the most dominant keywords, thus highlighting specific topics or issues of public concern regarding the DPR salary increase.

To understand the orientation of public opinion, a sentiment analysis was conducted using the *BERTweet – based sentiment* analysis model from Hugging Face. Each comment was classified into three sentiment categories: positive, negative, and neutral. This model was chosen for its ability to handle short and informal social media texts. The sentiment classification results provide an overview of the distribution of public opinion on the DPR salary issue, including tendencies toward support, criticism, or neutrality.

3. Results and Discussion

3.1. Sentiment Analysis Results

The sentiment distribution was then visualized in a bar chart to facilitate data interpretation. Each sentiment category is displayed along with the number of comments included in it. This visualization helps to see the overall proportion of positive, negative, and neutral opinions. Thus, this study not only provides quantitative information but also makes it easier for readers to visually understand public opinion trends. Based on the classification results using the Indonesian-RoBERTa Base Sentiment Classifier, the sentiment distribution is shown in Figure 1. Of the total comments analyzed, the majority of public responses indicated negative sentiment, amounting to 5,463 comments. Meanwhile, comments with neutral sentiment totaled 1,391, and only 812 positive sentiments. This significant dominance of negative sentiment indicates that the topic of the Dewan Perwakilan Rakyat (DPR) and the salaries of DPR members triggered a strong public reaction. The lack of positive sentiment indicates that this issue tends to generate criticism rather than appreciation.

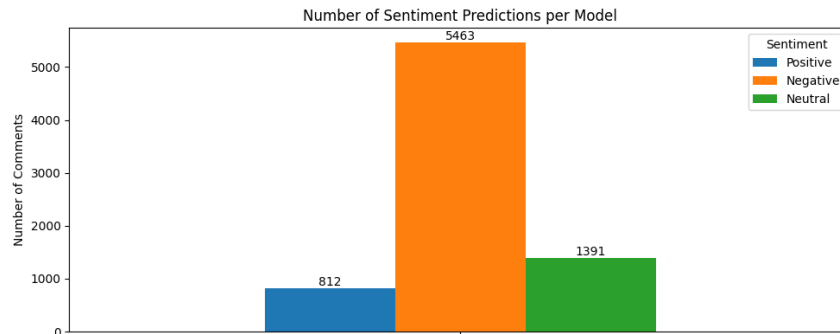


Figure 1: Sentiment Distribution Predicted by the Model

3.2. N-gram Analysis

To understand the context of the language used by the public, an n-gram analysis was conducted on YouTube comments. Table 1 below presents the frequency of the most frequently occurring words and phrases.

Table 1: Frequency of most frequently occurring words and phrases

Ngram Types	Ngram	Frequency
Unigram	<i>rakyat</i> (people / citizens)	4447
	<i>dpr</i> (House of Representatives)	3798
	<i>di</i> (in / at / on)	2210
	<i>yg</i> (which / that)	2123
	<i>gaji</i> (salary)	1516
Bigram	<i>anggota dpr</i> (members of the House of Representatives)	667
	<i>wakil rakyat</i> (representatives of the people)	401
	<i>rakyat jelata</i> (common people / ordinary citizens)	312
	<i>bubarkan dpr</i> (dissolve the House of Representatives)	305
	<i>rakyat yg</i> (people who / citizens that)	209
Trigram	<i>di atas penderitaan</i> (on top of suffering)	67
	<i>jadi anggota dpr</i> (become a member of the House of Representatives)	66
	<i>atas penderitaan rakyat</i> (upon the suffering of the people)	63
	<i>dias atas penderitaan rakyat</i> (on the suffering of the people)	56
	<i>bubarkan saja dpr</i> (just dissolve the House of Representatives)	44
4-gram	<i>di atas penderitaan rakyat</i> (on top of the people's suffering)	60
	<i>memang babie kalian semua</i> (you really are all babie [slang/insult])	24
	<i>ketua dpr ri rp</i> (Chairman of the Indonesian House of Representatives IDR)	16
	<i>menari diatas penderitaan rakyat</i> (dancing on the people's suffering)	15
	<i>menari di atas penderitaan</i> (dancing on suffering)	14

Table 1 shows that the most dominant unigrams are the words "rakyat" (4,447 times) and "DPR" (3,798 times). This confirms that public discourse is heavily focused on the relationship between the people as represented and the DPR as the legislative body. The word "salary," which appeared 1,516 times, also indicates that economic issues were a key driver of discussion. At the bigram level, the phrases "DPR members" (667 times) and "people's representatives" (401 times) indicate direct discussion of the identity and function of DPR members. Meanwhile, the phrase "disband the DPR" (305 times) indicates extreme dissatisfaction, giving rise to discourse about delegitimizing the institution. Furthermore, trigrams and 4-grams display strong emotional expressions, such as "above the people's suffering" and "dancing on the people's suffering." These phrases contain sarcastic criticism that confirms the public perception that the DPR does not side with the public's interests.

3.3. Comment Content Context

Komentar	Sentimen
Matane	Positive
Kerja nya anggota juga ngapain gak ada guna nya...	Negative
Apakah sudah saat nya kita kudeta pemerintahan	Negative
Gaji Pokok DPR RI – Ketua DPR: Rp5.040.000 per bulan, Wakil Ketua: Rp4.620.000 per bulan...	Neutral
Kalau memang memang kalian mewakili rakyat..., kenapa hanya mementingkan diri sendiri...	Negative

Figure 2: Comments and Predicted Sentiment

Analysis of comment snippets shows a predominance of negative tones. For example, comments like "What are the members doing? It's useless" and "Is it time for us to overthrow the government" reflect public dissatisfaction with the DPR's function. Neutral comments tend to be factual information, such as details of DPR salaries ("The basic salary of the Speaker of the DPR RI is Rp ... per month"), while positive comments are very rare and usually not directly related to the issue, for example, simply greetings or expressions of appreciation.

3.4. Discussion

The results of this study indicate a significant predominance of negative sentiment regarding the issue of Dewan Perwakilan Rakyat salary increases. This finding aligns with the explanation in the introduction that policies related to legislative salaries are considered sensitive because they directly impact social justice and government accountability. When the Dewan Perwakilan Rakyat salary policy is deemed disproportionate to the socioeconomic conditions of the community, such as high poverty rates, inequality, and budget constraints in priority sectors like health and education, negative perceptions emerge, leading to a decline in political legitimacy (Marcinkowski et al., 2024).

The n-gram analysis in this study reinforces this picture. The emergence of phrases such as "disband the Dewan Perwakilan Rakyat," "above the people's suffering," and "dancing on the people's suffering" reflects public frustration and cynicism. This phenomenon aligns with the explanation in the introduction that social media, particularly YouTube, has now become an important platform for directly channeling public aspirations and complaints (Isibor et al., 2021; Swastiningsih et al., 2024). The sarcastic and emotional comments demonstrate that the public is not only expressing rational criticism but also constructing a narrative of delegitimization of the Dewan Perwakilan Rakyat (DPR) as a representative institution.

The lack of positive sentiment indicates that the issue of DPR salary increases has little public support. This is crucial because political legitimacy is closely linked to public trust. If public trust continues to weaken, the DPR's continued function as a representative body will be increasingly questioned. This situation emphasizes the urgency of transparency and accountability in all state budget management policies, particularly those related to the welfare of public officials.

From a methodological perspective, the use of the IndoBERTa model (an Indonesian-based RoBERTa) has proven effective in capturing the complex linguistic nuances in YouTube comments, which often mix formal, informal, and emotional language. This is consistent with previous findings that transformer-based NLP models are capable of providing more accurate public opinion mapping than conventional methods (OJADI et al., 2023; Alam et al., 2025). Thus, this study not only provides an empirical picture of public opinion on the DPR salary issue but also contributes to the development of digital social data analysis methods in Indonesia.

The results of this study confirm that the issue of Dewan Perwakilan Rakyat salary increases is one of the most sensitive political issues in the digital space. The predominantly negative public response indicates deep dissatisfaction and highlights the importance of the House's responsiveness to public aspirations. If not managed effectively, such negative perceptions could tarnish the House's image and undermine its political legitimacy in the digital age, as outlined in the introduction.

4. Conclusion

This study found that public discourse on YouTube regarding the issue of Dewan Perwakilan Rakyat (DPR) salary increases was dominated by negative sentiment, with a significantly higher number of comments than neutral or positive sentiments. An n-gram analysis revealed that the most frequently occurring words and phrases related to "people," "DPR," and "salary," as well as emotional expressions such as "disband the DPR" and "dancing on the people's suffering." This reflects deep public dissatisfaction with the DPR's salary policy, which is considered inconsistent with the socioeconomic conditions of the community.

These findings support the argument in the literature that issues related to the welfare of public officials, particularly legislative salaries, often trigger crises of trust and undermine political legitimacy. The lack of positive comments also indicates low public support, further questioning the legitimacy of the DPR as a representative institution.

From a methodological perspective, the use of the IndoBERTa model has proven capable of capturing linguistic nuances in public comments, including informal, mixed language, and emotional expressions. Thus, this study not only provides an empirical picture of digital public opinion on the DPR salary increase issue but also offers a methodological contribution to big data-based public opinion studies in Indonesia.

Overall, the results of this study emphasize the importance of transparency, accountability, and responsiveness by the DPR in all policies related to the welfare of public officials. Failure to do so could lead to widespread negative perceptions in the digital space that could damage the DPR's image and erode its political legitimacy in the eyes of the public.

References

- Alam, M. S., Mrida, M. S. H., & Rahman, M. A. (2025). Sentiment analysis in social media: How data science impacts public opinion knowledge integrates natural language processing (NLP) with artificial intelligence (AI). *American Journal of Scholarly Research and Innovation*, 4(01), 63-100.
- Babu, Y. P., & Eswari, R. (2021). Sentiment Analysis on Dravidian Code-Mixed YouTube Comments using Paraphrase IndoBERTaModel. In *FIRE (Working Notes)* (pp. 1069-1076).
- Isibor, N. J., Ewim, C. P. M., Ibeh, A. I., Adaga, E. M., Sam-Bulya, N. J., & Achumie, G. O. (2021). A generalizable social media utilization framework for entrepreneurs: Enhancing digital branding, customer engagement, and growth. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(1), 751-758.
- Klingenger, L., Shahi, S., Frère, E., & Zureck, A. (2022). Inclusive Measurement of Public Perception of Corporate Low-Carbon Ambitions: Analysis of Strategic Positioning for Sustainable Development Using Natural Language Processing. *International Journal of Sustainable Development & Planning*, 17(1).
- Marcinkowski, L., Butnaru, A., & Rabrenović, A. (2024). Salary systems in public administration and their reforms. *SIGMA Papers*.
- Ojadi, J. O., Onukwulu, E. C., Somtochukwu, C., & Odionu, O. A. O. (2023). Natural Language Processing for Climate Change Policy Analysis and Public Sentiment Prediction: A Data-Driven Approach to Sustainable Decision-Making. *Iconic Research and Engineering Journals*, 7(3), 732-751.
- Stefanovic, N., Radenkovic, M., Bogdanovic, Z., Plasic, J., & Gaborovic, A. (2025). Adaptive cloud-based big data analytics model for sustainable supply chain management. *Sustainability*, 17(1), 354.
- Swastiningsih, S., Aziz, A., & Dharta, Y. (2024). The Role of Social Media in Shaping Public Opinion: A Comparative Analysis of Traditional vs. Digital Media Platforms. *The Journal of Academic Science*, 1(6), 620-626.