# Sentiment Analysis of Public Opinions Regarding "Ideas of Presidential Candidates" in YouTube Video Comments with Robustly Optimized BERT Pretraining Approach

*Yo'el Pieter Sumihar*
*Universitas Kristen Immanuel*
*Email : pieter.haro@ukrimuniversity.ac.id*

## ABSTRACT

Social media and video-sharing platforms such as YouTube have become one of the primary sources of information and social interaction in modern society. In politics, YouTube has become essential for spreading ideas, campaign platforms, and opinions about the presidential election. This study aims to analyze public sentiment towards "Presidential Candidate Ideas" conveyed in videos on the YouTube platform. It can be a benchmark for comparing sentiment between potential presidential candidates. The research began by collecting data from YouTube comments, preprocessing it to remove unnecessary data, and testing it using the pre-trained Indonesian Roberta Base Sentiment Classifier Model to obtain sentiment results. Using the pre-trained Indonesian Roberta Base Sentiment Classifier Model, the data obtained from YouTube comments will be divided into positive, negative, and neutral labels. The results of this study are the accuracy for each sentiment label, where the value for positive is 93%, the negative is 90.5%, and the neutral is 93.04%. Residents comment more positively on presidential candidate Prabowo Subianto, with a positive value of 54.13%, followed by Anies Baswedan at 42.8% and Ganjar Pranowo at 31.91%.

**Keywords:** nlp, Roberta, youtube, comment, sentiment.

## INTRODUCTION

Social media and video-sharing platforms such as YouTube have become one of the primary sources of information and social interaction in modern society (Vira & Reynata, 2022). In a political context, YouTube has become an important platform for disseminating ideas, campaign platforms, and opinions regarding the presidential election. In facing the presidential election, the ideas and opinions expressed by presidential candidates or potential presidential candidates are fundamental to understanding people's preferences and supporting a democratic decision-making process.

Public sentiment toward presidential candidates' ideas is very influential in the election process. Therefore, accurate sentiment analysis of public opinion on YouTube media can provide valuable insights to political stakeholders, researchers, and the general public.

The importance of sentiment analysis in the context of presidential elections has been uncovered in previous research. The application of sentiment analysis, particularly in social media, has experienced a rapid and dynamic growth in recent years. One method that has garnered significant attention in sentiment analysis is the use of Transformer-based language models, such as BERT (Bidirectional Encoder Representations from Transformers).

The use of the RoBERTa (Robustly Optimized BERT Pretraining) approach has emerged as a trend in recent research (Zain et al., 2021). RoBERTa, a variation of the BERT model, has been optimized with larger and better training methods. This method has demonstrated significant success in a variety of natural language processing applications, instilling confidence in its potential for sentiment analysis.

Despite advancements in sentiment analysis and models like RoBERTa, a gap exists in understanding public sentiment toward presidential candidates on YouTube. While most studies focus on text-based platforms like Twitter and Facebook, less attention has been given to video-based platforms where content and user engagement differ. This study addresses these gaps by analyzing YouTube comments on presidential candidates using advanced Transformer-based models to capture evolving public sentiment.

## LITERATURE REVIEW

Sentiment Analysis is a method for extracting information from text to identify sentiments that can be classified as positive, neutral, or negative. As the presidential election approaches, there is a lot of news and views circulating on social media regarding the candidates. These candidates can use this information to understand positive public views as potential political support that needs to be strengthened and opposing views that need to be corrected. To map these views, a text-based opinion classification system is needed. Opinions will be grouped into three categories: negative, neutral, and positive. This action will make it easier for presidential candidates to find out the sentiments of the public regarding the ideas that have been conveyed in the YouTube video.

The first study by Zain and colleagues in 2021 (Zain et al., 2021) has implemented the Robustly Optimized BERT Pretraining approach in analyzing public opinion sentiment regarding the

COVID-19 vaccine on Twitter social media. They used the pre-trained Indonesian RoBERTa Base Sentiment Classifier to classify data into three different classes, namely positive, neutral and negative. This study's results show impressive accuracy levels, with average prediction accuracy reaching 84% for positive labels, 97% for neutral labels, and 93% for negative labels.

The second study, conducted by Farah Zhafira and the team in 2021 (Farah Zhafira et al., 2021), focused on sentiment analysis regarding the Independent Campus Policy using comment data on the YouTube platform. They applied positive and negative sentiment classification methods with Naive Bayes and Tf-Idf weighting. The results of this study show a high level of accuracy, with an average accuracy reaching 91.8% through 10 iterations of k-fold cross-validation, as well as precision, recall, and f-measure values reaching 90.35%, 93.6%, and 91.95%.

The third study conducted by Daffa and colleagues in 2023 (Daffa et al., 2023) explored public sentiment towards the Corporate Social Responsibility (CSR) program at PT. Olahkarsa by using the SR-App Olahkarsa application. The results of this study reflect very positive reception from users, with a positive acceptance rate reaching 95%.

The three studies show significant developments in public sentiment analysis on various topics using various methods and platforms. They provide important insights into how society reacts to contemporary issues through data obtained from social media and video sharing platforms.

## METHODS

The research methods that will be carried out in this research can be seen in Figure 1 :

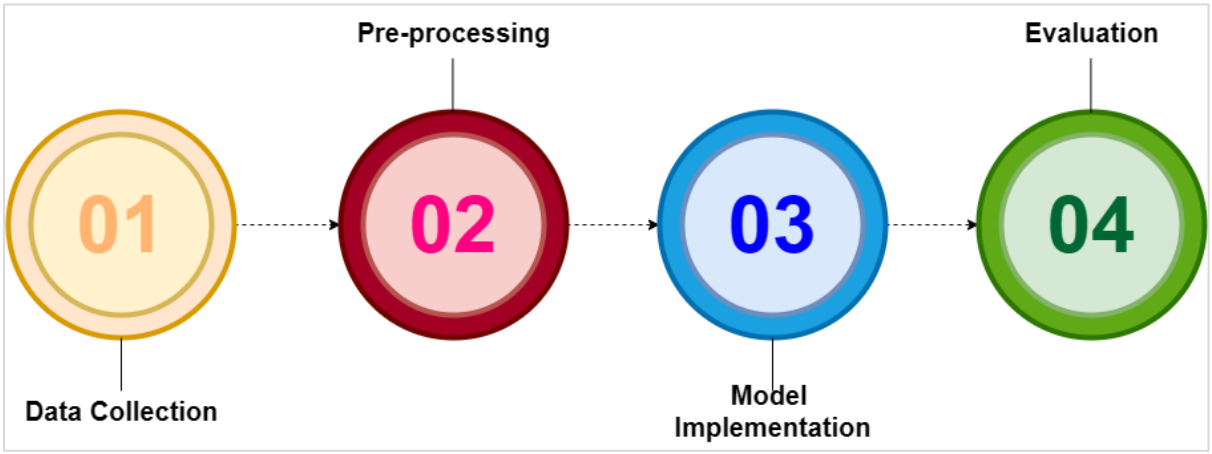


**Figure 1.** Research Methods

**Data collection**

The technique used to collect data is to use the YouTube API, which Google API provides. Google API Client is a tool that can be used to access many things related to Google. One of the libraries owned by Google API Client is the YouTube Data API. This YouTube API data library has many uses, including retrieving comments on specific videos for free.

In this research, the video used to collect data is a video on Najwa Sihab's YouTube channel talking to the three presidential candidates for the 2024 period, which has been divided into three different videos with the following links:

1. Anies Baswedan Talks Ideas - https://www.youtube.com/watch?v=kiaKPHMABuc
2. Ganjar Pranowo Talks Ideas - https://www.youtube.com/watch?v=2YXKMHNevpo
3. Prabowo Subianto Talks Ideas - https://www.youtube.com/watch?v=V4W5Nokc7MU

Comments will be taken from each video to be saved, and the sentiment value will be tested later.

**Text Pre-processing**

The data pre-processing stage is crucial in the text data classification process. The data pre-processing process aims to remove noise or irrelevant information, such as emojis, URLs, and numbers, and to produce a uniform word format so that the word data becomes cleaner before the next step. In the data obtained, words that do not have a clear structure are often found. Therefore, data processing or cleaning actions, such as case folding, are needed to change all letters to lowercase. After that, the data cleaning process is carried out, which involves removing duplicate data, removing symbols, URLs, numbers, ASCII characters, emojis, abbreviations such as 'yg,' 'dg,' 'tidak,' and punctuation such as ["™," "©,""℠"].

**Indonesian RoBERTa Base Sentiment Classifier**

After the data has been cleaned, the next step is to perform text processing and labeling using a pre-trained model, namely the Indonesian Roberta Base Sentiment Classifier model (Wongso, 2023). The Indonesian sentiment classification based on RoBERTa is a model used to classify sentiment in text based on the RoBERTa model. Initially, this model was a previously trained Indonesian RoBERTa Base model, then improved using the Indonesian SmSA dataset

consisting of Indonesian comments and reviews. After the training process, this model achieved an accuracy level of 94.36% and an F1-macro of 92.42%. When tested on the benchmark dataset, this model achieved an accuracy of 93.2% and an F1-macro of 91.02% (Wongso, 2023).

## RESULTS

### Scraping

The results found from the web scraping technique on each video amounted to 1500 comments on each video. The video's comments were retrieved using the API that Google officially provided using Python. After the scraping process was carried out, then this raw data was stored, and five attributes were obtained consisting of the author (user), published_at (time the comment was made), updated_at (time the comment was updated), like_count (number of likes on the comment), text (raw comment data). However, from these attributes, the one that has an important role in text mining is the "text" attribute, which is data from YouTube user comments that, if processed, will produce much knowledge. After the data scraping activity, the number of tweets ready for the Pre-processing process and analysis was 1500 data for each video representing each presidential candidate. Details of the amount of data can be seen in table 1, and examples of comments on each video can be seen in table 2.

**Table 1**. Number of comments for each video

| No | Video ID | Video Title | Comment Count |
|---|---|---|---|
| 1. | kiaKPHMABuc | Anies Baswedan Talks Ideas | 1549 |
| 2. | V4W5Nokc7MU | Prabowo Subianto Talks Ideas | 1553 |
| 3. | 2YXKMHNevpo | Ganjar Pranowo Talks Ideas | 1571 |

**Table 2**. Example comment for each video

| No | Video ID | Comment |
|---|---|---|
| 1. | kiaKPHMABuc | Klo bicara gagasan dan wawasan Pak Anies lebih dapet sih. Lugas, cerdas, realistis dan logis ❤❤❤❤❤ |
| | | Rakyat indonesia butuh karya nyata jgn asal narasi siapapun dia yg terpilih |
| | | Siapa pun pemimpin nya.. yang penting jujur dan bijak didalam segala Hal/bidang |
| 2. | V4W5Nokc7MU | All in Prabowo, beliau ketua partai jadi gak akan ada yang nyetir, bismillah 2024 Prabowo Subianto RI 1 |

| | | |
|---|---|---|
| | | Ya allah jadikanlah bpk prabowo ini pemimpin negara Indonesia ya allah aminn. |
| | | Pak Prabowo cuma nggak pinter ngomong. Tetapi pemikirannya strategis, wataknya jujur, jiwanya patriot, hatinya baik, tujuannya lurus. |
| 3. | 2YXKMHNevpo | awalnya saya meragukan pak Ganjar, tapi setelah mendengar jawabannya saya semakin ragu :\| |
| | | Sayang sekali pak Ganjar TDK mengetahui kemampuan anak bangsa untuk mengelola perusahaan yg berteknologi tinggi. |
| | | Seneng bgt jika penonton vidio ini lebih dari 100 jt, artinya banyak orang yang mulai suka dan peduli dengan bangsa ini |

**Pre-processing**

The pre-processing or data preparation stage is the most crucial stage in a study that uses text mining. Raw data from the results of text mining activities must be prepared in advance according to the analysis needs so that it can produce good analysis results. For example, in this study, one of the stages in the pre-processing activity is changing all font sizes to lowercase. This is important because the system will read a word consisting of the same letters as a different word if the letters have different letter sizes. For example, the words 'Ceria' and 'ceria'. If this is left alone, it will affect the computer's time to process the data. However, most importantly, the results obtained will be ambiguous or unclear.

In this study, the author manages the data obtained through previous scraping, with examples of differences in data before and after pre-processing as follows:

**Table 3**. Example comparison of original comments and preprocessing results

| No | Original Comment | Preprocessing Results |
|---|---|---|
| 1. | Klo bicara gagasan dan wawasan Pak Anies lebih dapet sih. Lugas, cerdas, realistis dan logis ❤❤❤❤❤ | klo bicara gagasan dan wawasan pak anies lebih dapet sih. lugas, cerdas, realistis dan logis |
| | Rakyat indonesia butuh karya nyata jgn asal narasi siapapun dia yg terpilih | rakyat indonesia butuh karya nyata jgn asal narasi siapapun dia yang terpilih |
| | Siapa pun pemimpin nya.. yang penting jujur dan bijak didalam segala Hal/bidang | siapa pun pemimpin nya.. yang penting jujur dan bijak didalam segala hal bidang |
| 2. | All in Prabowo, beliau ketua partai jadi gak akan ada yang nyetir, bismillah 2024 Prabowo Subianto RI 1 | all in prabowo, beliau ketua partai jadi gak akan ada yang nyetir, bismillah 2024 prabowo subianto ri 1 |
| | Ya allah jadikanlah bpk prabowo ini pemimpin negara Indonesia ya allah aminn. | ya allah jadikanlah bpk prabowo ini pemimpin negara indonesia ya allah aminn. |

| | | |
|---|---|---|
| | Pak Prabowo cuma nggak pinter ngomong. Tetapi pemikirannya strategis, wataknya jujur, jiwanya patriot, hatinya baik, tujuannya lurus. | pak prabowo cuma nggak pinter ngomong. tetapi pemikirannya strategis, wataknya jujur, jiwanya patriot, hatinya baik, tujuannya lurus. |
| 3. | awalnya saya meragukan pak Ganjar, tapi setelah mendengar jawabannya saya semakin ragu :\| | awalnya saya meragukan pak ganjar, tapi setelah mendengar jawabannya saya semakin ragu |
| | Sayang sekali pak Ganjar TDK mengetahui kemampuan anak bangsa untuk mengelola perusahaan yg berteknologi tinggi. | sayang sekali pak ganjar tdk mengetahui kemampuan anak bangsa untuk mengelola perusahaan yg berteknologi tinggi. |
| | Seneng bgt jika penonton vidio ini lebih dari 100 jt, artinya banyak orang yang mulai suka dan peduli dengan bangsa ini | seneng bgt jika penonton vidio ini lebih dari 100 jt, artinya banyak orang yang mulai suka dan peduli dengan bangsa ini |

**Indonesian RoBERTa Base Sentiment Classifier**

The results of the cleaned data will be implemented into the pre-trained Indonesian RoBERTa Base Sentiment Classifier model to produce labels and ensure the accuracy of label predictions. Below are the results of the implementation and the implementation of the Indonesian RoBERTa Base Sentiment Classifier. The prediction process takes approximately 1 hour and 12 minutes to classify sentiment on 4500 comments.

**Table 4**. Example of Indonesian RoBERTa Base Sentiment Classifier execution result

| No | Comment | Positif | Netral | Negatif | Label |
|---|---|---|---|---|---|
| 1. | ayo selamatkan bangsa ini negri ini dengan tidak memilih anggota pdip atau yang berhubungan dengan pdip | 0.79 | 98.86 | 0.35 | Netral |
| 2. | awalnya saya meragukan pak ganjar tapi setelah melihat semua jawaban dan sikap beliau ketika menjawab pertanyaan di video ini, saya menjadi yakin, yakin tidak akan memilih beliau | 1.09 | 98.3 | 0.6 | Netral |
| 3. | pak prabowo memang sulit untuk merangkai kata kata dalam berbicara. tapi soal kecerdasan. aksi nya. pola pikirnya. beliau yang paling unggul. | 99.84 | 0.11 | 0.04 | Positif |

In the example above, you can see that there are positive, neutral, and negative columns. The columns contain the percentage value of trust in sentiment, and the results are stored in the label column.

After sentiment checking was carried out on all data using the data model, the average accuracy results of the pre-trained model were obtained, which can be seen in Table 5.

**Table 4**. Average accuracy results

| No | Label | Average Accuracy |
|---|---|---|
| 1. | Positif | 93,00% |
| 2. | Netral | 93,04% |
| 3. | Negatif | 90,52% |

From these results, it can be seen that the average accuracy of the sentiment model obtained results above 90%, which proves that the model is very good for identifying sentiment in Indonesian.

**Visualization of Anies Baswedan's Comment Analysis Results**

The sentiment value for each comment in Anies Baswedan’s ideas video was obtained from a pre-trained Indonesian language sentiment model. It can be seen from the following table that the response from the majority of YouTube viewers of Anies Baswedan's Idea is Positive at 42.8%, Negative at 20.4%, and Neutral at 36.8%. These results indicate that many viewers like Anies Baswedan as a presidential candidate, but the positive value is not large enough because it is not far from the neutral value. From these results, it can also be seen that Anies' Idea is quite good because the negative value is relatively low.

The following is a table of the number and percentage of public responses to Anies Baswedan's ideas based on the type of sentiment:

**Table 5**. Responses To Anies Baswedan's Ideas Based

| No | Label | Amount | Percentage |
|---|---|---|---|
| 1. | Positif | 653 | 42,79% |
| 2. | Netral | 561 | 36,76% |
| 3. | Negatif | 312 | 20,44% |

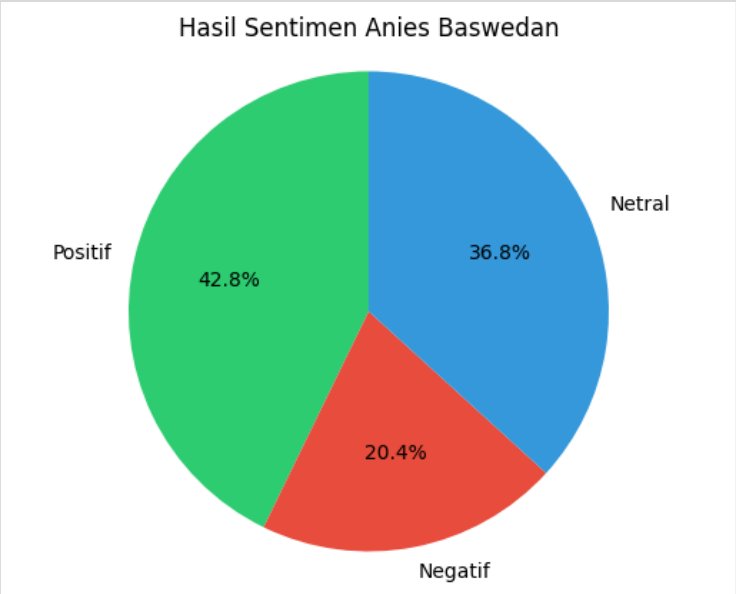


**Figure 2**. Responses To Anies Baswedan's Ideas Based

The following is a graph of the frequency of sentiment regarding public responses to Anies Baswedan's idea as a presidential candidate from day to day according to the comment date on the YouTube video:

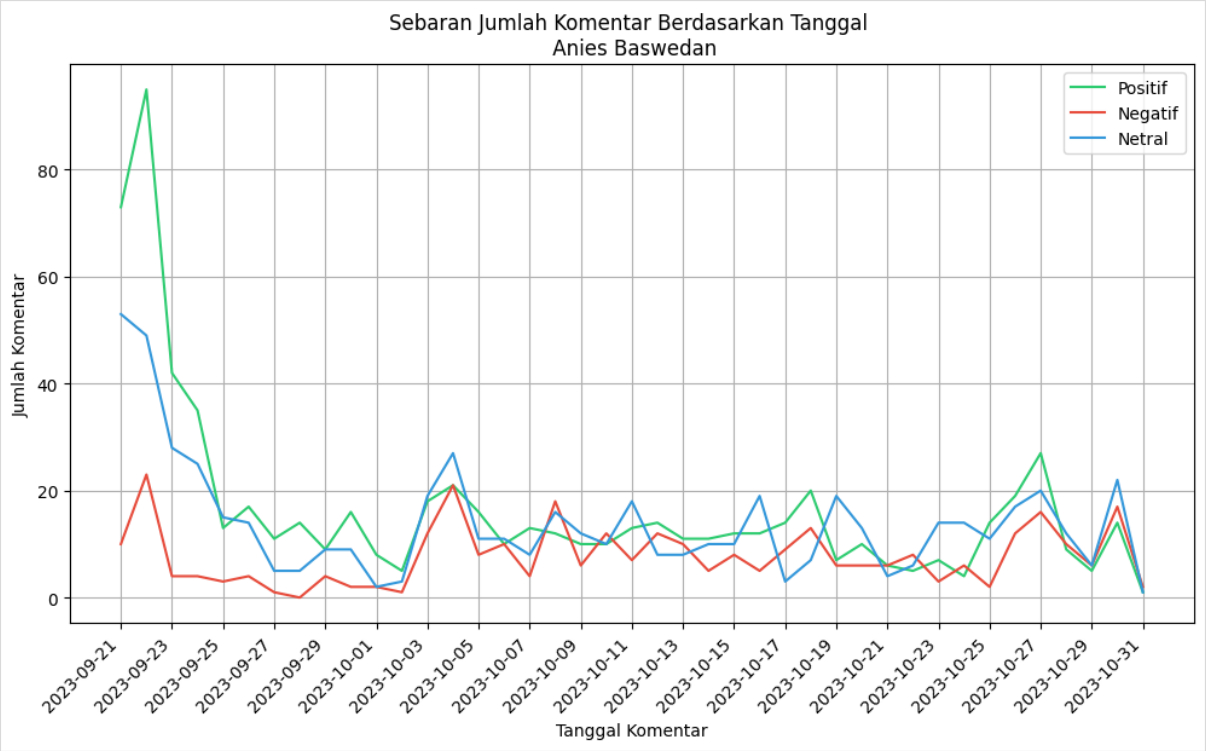


**Figure 3**. Responses To Anies Baswedan's Ideas from day to day

In Figure 3, it can be seen that at the beginning of the video's publication, there were more positive comments than neutral and negative comments. There was also an increase in positive comments the next day. However, it can be seen that day by day, both positive, negative, and neutral comments have an average comment that is not too different.

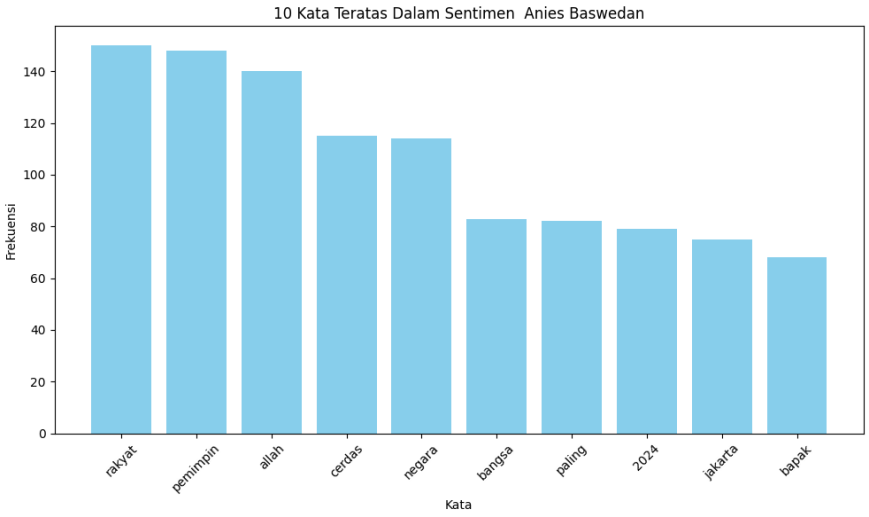


**Figure 4.** Top 10 Words in Anies Baswedan's Sentiment


**Figure 5.** Word Cloud Analysis in Anies Baswedan's Sentiment

Figure 4 shows the ten most frequently appearing words in YouTube comments for Anies Baswedan, and Figure 5 shows the most frequently appearing words in all YouTube comments for Anies Baswedan.

### Visualization of Prabowo Subanto's Comment Analysis Results

The sentiment value for each comment in Prabowo Subianto’s ideas video was obtained from a pre-trained Indonesian language sentiment model. It can be seen from the following table that

the response from the majority of YouTube viewers of Prabowo Subianto's Idea is Positive at 54,13%, Negative at 16,5%, and Neutral at 29,3%. These results identify that more than 50% of the audience likes the idea of Prabowo as a presidential candidate, and this can also be compared with the low negative sentiment and neutral sentiment values which further confirm that the audience likes or agrees with the idea of Prabowo as a presidential candidate.

The following is a table of the number and percentage of public responses to Prabowo Subianto's ideas based on the type of sentiment:

**Table 6**. Responses To Prabowo Subianto's Ideas Based

| No | Label | Amount | Percentage |
|---|---|---|---|
| 1. | Positif | 825 | 54,13% |
| 2. | Netral | 447 | 29,33% |
| 3. | Negatif | 252 | 16,53% |

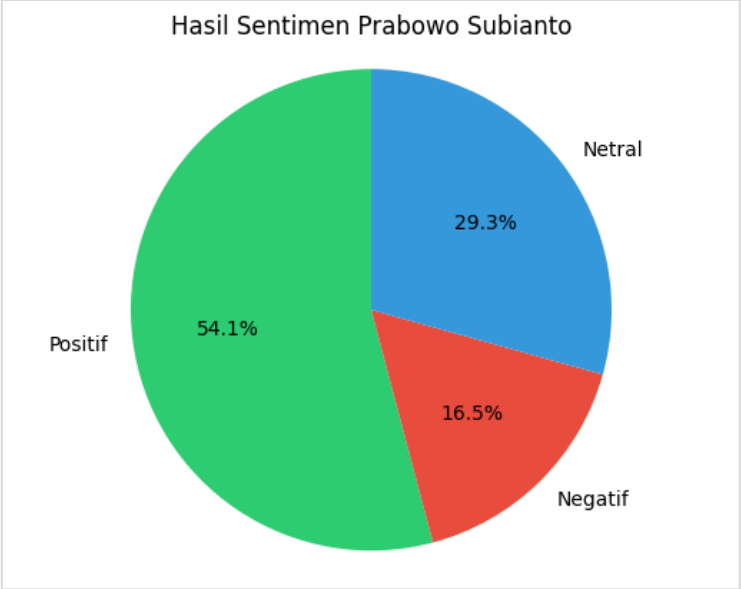


**Figure 6**. Responses To Prabowo Subianto's Ideas Based

The following is a graph of the frequency of sentiment regarding public responses to Prabowo Subianto's idea as a presidential candidate from day to day according to the comment date on the YouTube video:

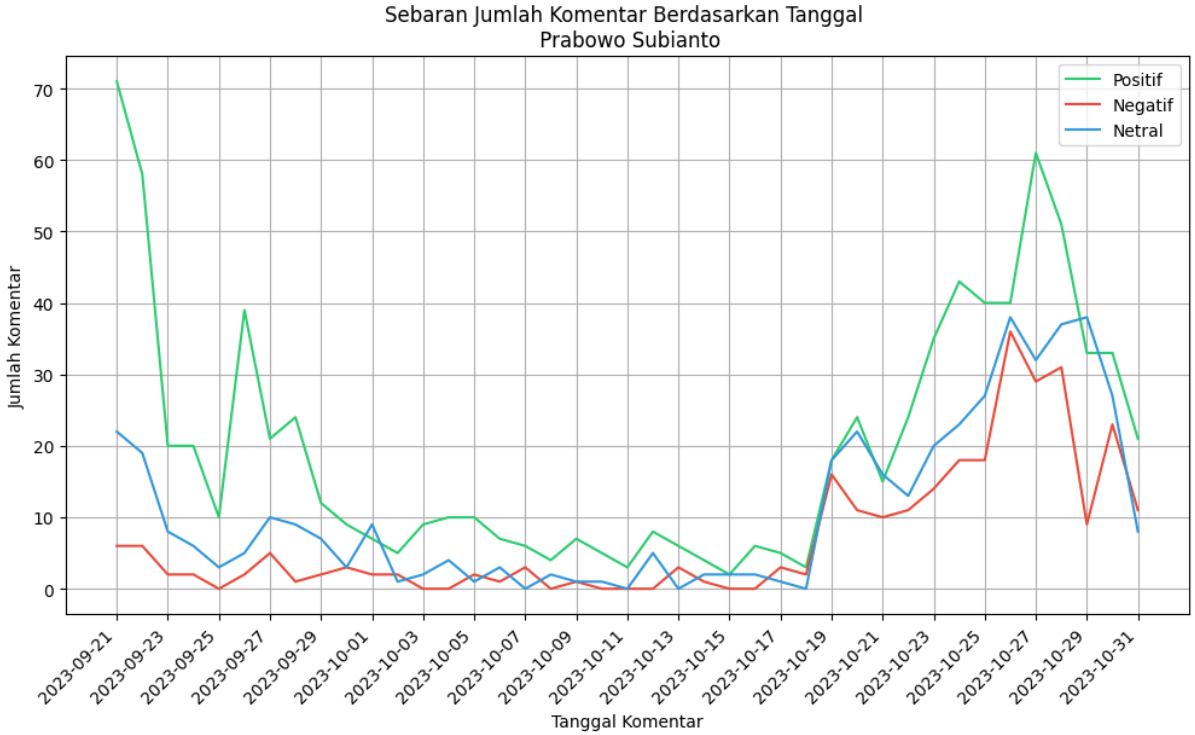


**Figure 7**. Responses To Prabowo Subianto's Ideas from day to day

In Figure 7, it can be seen that at the beginning of the video's publication, there were more positive comments than neutral and negative comments. There was also an increase in positive comments the next day. However, it can be seen that day by day, both positive, negative, and neutral comments have an average comment that is not too different. But in the last week, there is an increase for positive, negative and netral comment.

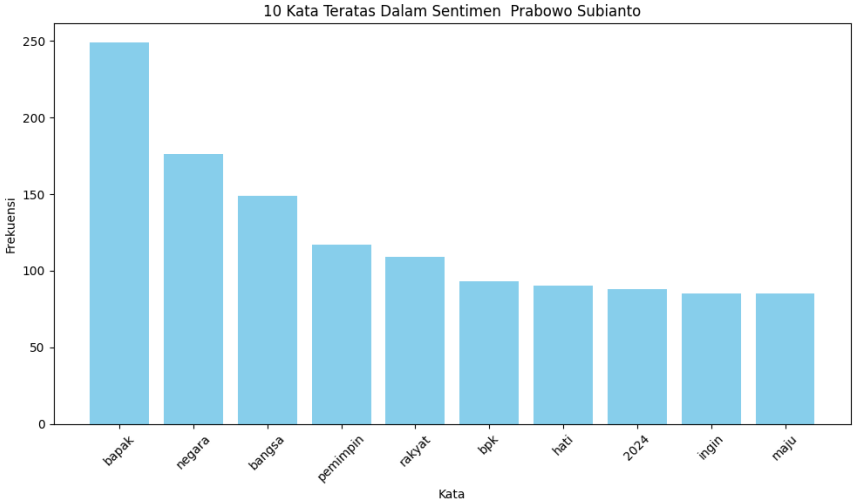


**Figure 8.** Top 10 Words in Prabowo Subianto's Sentiment

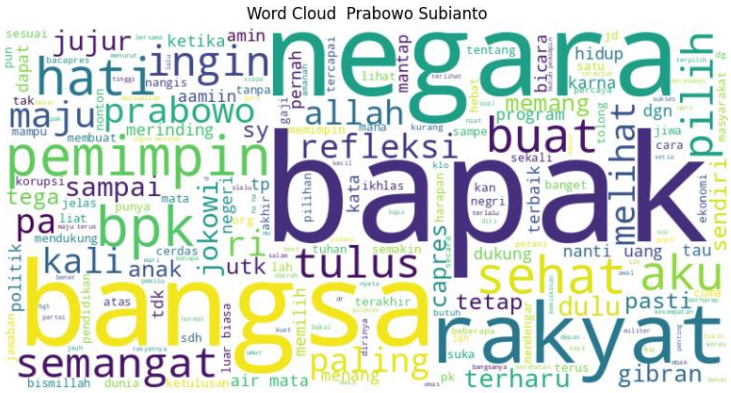


**Figure 9.** Word Cloud Analysis in Prabowo Subianto's Sentiment

Figure 8 shows the ten most frequently appearing words in YouTube comments for Probowo Subianto, and Figure 9 shows the most frequently appearing words in all YouTube comments for Prabowo Subianto.

**Visualization of Ganjar Pranowo's Comment Analysis Results**

The sentiment value for each comment in Ganjar Pranowo’s ideas video was obtained from a pre-trained Indonesian language sentiment model. It can be seen from the following table that the response from the majority of YouTube viewers Ganjar Pranowo’s Idea is Netral at 43,34%, Negative at 24,73%, and Positive at 31,91%. This result identifies that Ganjar's idea received only a few positive comments from citizens by getting a value not too far from the neutral comment. However, it does not mean that Ganjar Pranowo's idea is terrible because the negative value obtained is relatively low.

The following is a table of the number and percentage of public responses to Ganjar Pranowo’s ideas based on the type of sentiment:

**Table 7**. Responses To Ganjar Pranowo’s Ideas Based

| No | Label | Amount | Percentage |
|---|---|---|---|
| 1. | Positif | 489 | 31,91% |
| 2. | Netral | 664 | 43,34% |
| 3. | Negatif | 379 | 24,73% |

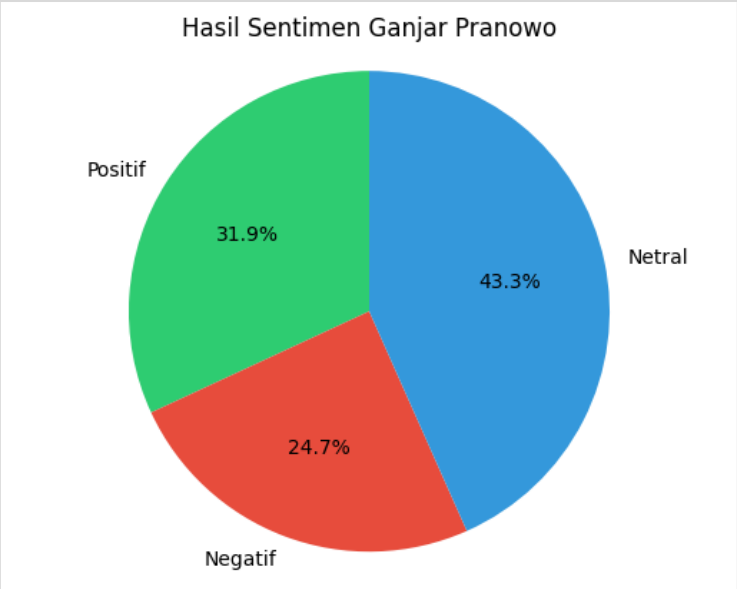


**Figure 10**. Responses To Ganjar Pranowo's Ideas Based

The following is a graph of the frequency of sentiment regarding public responses to Ganjar Pranowo's idea as a presidential candidate from day to day according to the comment date on the YouTube video:

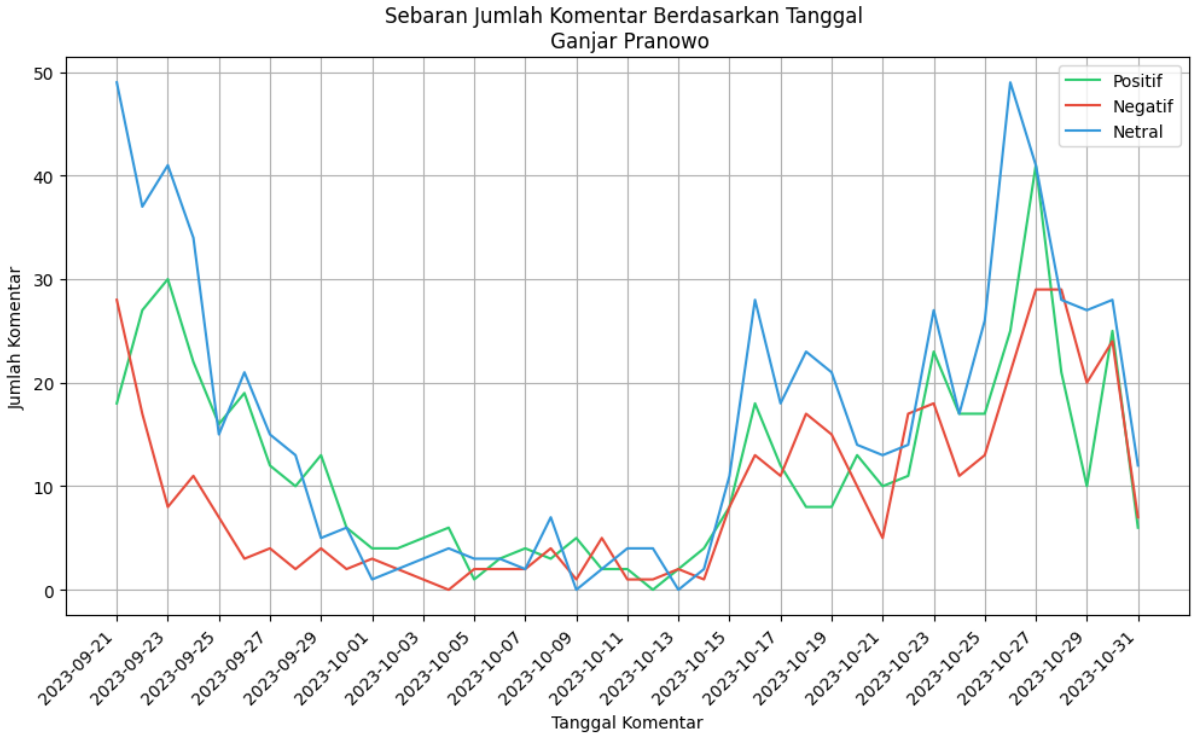


**Figure 11**. Responses To Ganjar Pranowo's Ideas from day to day

In Figure 11, it can be seen that at the beginning of the video's publication, there were more neutral comments than positive and negative comments. However, it can be seen that day by day, both positive, negative, and neutral comments have an average comment that is not too different. But in the last week, there is an increase for positive, negative and netral comment.

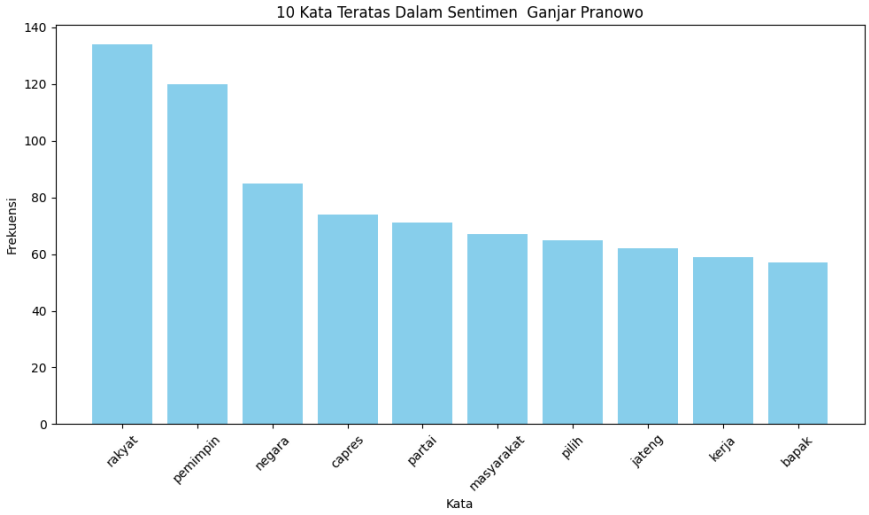


**Figure 12.** Top 10 Words in Ganjar Prabowo's Sentiment

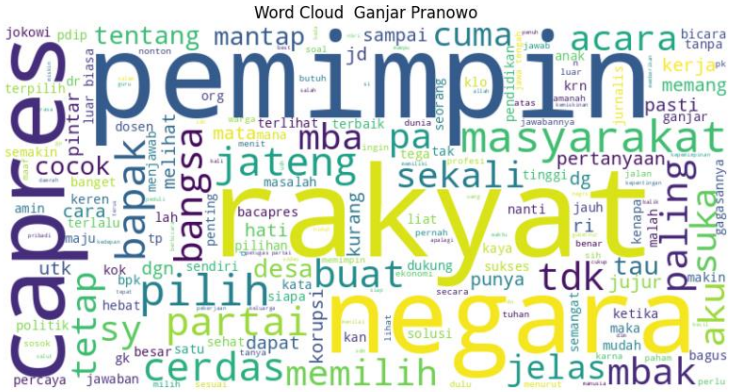


**Figure 13.** Word Cloud Analysis in Ganjar Prabowo's Sentiment

Figure 12 shows the ten most frequently appearing words in YouTube comments for Ganjar Prabowo, and Figure 13 shows the most frequently appearing words in all YouTube comments for Ganjar Pranowo.

## CONCLUSION

A summary of sentiment analysis research by implementing a pre-trained Indonesian RoBERTa Base Sentiment Classifier, where data is classified into 3 different classes, has been conducted. Based on the results of the research, it can be concluded that the accuracy for each sentiment label is good, where the value for positive is 93%, negative is 90.5%, and neutral is

93.04%. Based on the data obtained, it is also concluded that residents gave more positive comments to presidential candidate Prabowo Subianto with a positive value of 54.13% followed by Anies Baswedan at 42.8% and Ganjar Pranowo at 31.91%.

## REFERENCES

Daffa, M., Rifqi, A., & Rizky Yunianto, D. (2023). ANALISIS SENTIMEN BERITA PROGRAM CSR PADA APLIKASI SR-APP OLAHKARSA. Jurnal Informatika dan Teknik Elektro Terapan, 11(3), 2830–7062. https://doi.org/10.23960/jitet.v11i3%20s1.3413

Farah Zhafira, D., Rahayudi, B., & Korespondensi, P. (2021). ANALISIS SENTIMEN KEBIJAKAN KAMPUS MERDEKA MENGGUNAKAN NAIVE BAYES DAN PEMBOBOTAN TF-IDF BERDASARKAN KOMENTAR PADA YOUTUBE (Vol. 2, Nomor 1).

Laia, Y.; Berutu, S.; Sumihar, Y.; Budiati, H. Implementasi Library Textblob Dan Metode Support Vector Machine Pada Analisis Sentimen Pelanggan Terhadap Jasa Transportasi Online. bits 2024, 6, 1−10.

Lunando, E., & Purwarianti, A. (2013). *Indonesian Social Media Sentiment Analysis with Sarcasm Detection*.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. http://arxiv.org/abs/1907.11692

Multi Fani, S., & Santoso, R. (2021). *PENERAPAN TEXT MINING UNTUK MELAKUKAN CLUSTERING DATA TWEET AKUN BLIBLI PADA MEDIA SOSIAL TWITTER MENGGUNAKAN K-MEANS CLUSTERING*. *10*, 583–593. https://ejournal3.undip.ac.id/index.php/gaussian/

Matresya Matulatuwa, F., Sediyono, E., & Iriani, A. (2017). *TEXT MINING DENGAN METODE LEXICON BASED UNTUK SENTIMENT ANALYSIS PELAYANAN PT. POS INDONESIA MELALUI MEDIA SOSIAL TWITTER* (Vol. 2, Issue 3).

Sihab, N. (n.d.). Anies Baswedan Bicara Gagasan. Retrieved September 27, 2023, from https://www.youtube.com/watch?v=kiaKPHMABuc

Sihab, N. (n.d.). Prabowo Subianto Bicara Gagasan. Retrieved September 27, 2023, from https://www.youtube.com/watch?v= V4W5Nokc7MU

Sihab, N. (n.d.). Ganjar Pranowo Bicara Gagasan. Retrieved September 27, 2023, from https://www.youtube.com/watch?v=2YXKMHNevpo

Vira, A., & Reynata, E. (2022). PENERAPAN YOUTUBE SEBAGAI MEDIA BARU DALAM KOMUNIKASI MASSA. Komunikologi : Jurnal Ilmiah Ilmu Komunikasi, 19(2), 96–101.

Wongso, W. (2023). Indonesian RoBERTa Base Sentiment Classifier. https://huggingface.co/w11wo/indonesian-roberta-base-sentiment-classifier. https://huggingface.co/w11wo/indonesian-roberta-base-sentiment-classifier

Yunitasari, Y., Musdholifah, A., & Sari, A. K. (2019). Sarcasm Detection For Sentiment Analysis in Indonesian Tweets. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *13*(1), 53. https://doi.org/10.22146/ijccs.41136

Zain, M. M., Simbolon, R. N., Sulung, H., & Anwar, D. Z. (2021). Jurnal Politeknik Caltex Riau Analisis Sentimen Pendapat Masyarakat Mengenai Vaksin Covid-19 Pada Media Sosial Twitter dengan Robustly Optimized BERT Pretraining Approach. Dalam Jurnal Komputer Terapan (Vol. 7, Nomor 2). https://jurnal.pcr.ac.id/index.php/jkt/

Zendrato, A.; Berutu, S.; Sumihar, Y. P.; Budiati, H. Pengembangan Model Klasifikasi Sentimen Dengan Pendekatan Vader Dan Algoritma Naive Bayes Terhadap Ulasan Aplikasi Indodax. josh 2024, 5, 755-764.