

# Sentiment Analysis Of Twitter Towards The 2024 Indonesian Presidential Candidates Using The Naïve Bayes Algorithms

Yogi Gunawan<sup>1\*</sup>, Iwan Purnama<sup>2</sup>, Rohani<sup>3</sup>

<sup>1,2,3</sup> Faculty of Science and Technology, Universitas Labuhanbatu, Sumatera Utara Indonesia.

\*Corresponding Author:

Email: [yogigunawan173@gmail.com](mailto:yogigunawan173@gmail.com)

---

## Abstract.

*The increasing use of social media (Twitter) has made it a platform for the public to express their views on the Indonesian presidential candidate in the 2024 elections. The sentiment expressed through comments on Twitter provides important insights into the public perception of the candidates. However, given the volume and speed at which information is disseminated on social media, manual analysis of this sentiment becomes impractical. Therefore, the use of the Naïve Bayes algorithm for automatic sentiment analysis is considered essential to understanding voter support and preferences. The study aims to analyze Twitter users' sentiments towards three Indonesian presidential candidates in 2024, Anies, Ganjar, and Prabowo, using the Naïve Bayes algorithm. We categorize the results of this analysis into three sentiment categories: positive, negative, and neutral. The methods used in the study involved collecting Twitter comment data related to the three candidates, pre-processing data, labeling data, applying the Naïve Bayes algorithm for the classification of sentiment, and evaluation of the performance of the algorithm performed by calculating the level of accuracy. The results of the research showed that the Naïve Bayes algorithm was able to classify sentiments with fairly high precision, namely 75.54% for Anies, 82.74% for Ganjar, and 75.24% for Prabowo. The conclusion of this study is that sentimental analysis using the Naïve Bayes algorithm can provide significant insights into voter preferences and support. The sentimental data generated can serve as a strong foundation for decision-makers to design campaign strategies that are more effective and responsive to public perception. This research also opens up opportunities for further development in the use of sentimental analysis techniques in politics and campaigns.*

**Keywords:** Elections, Presidential Candidates, Naïve Bayes, Sentiment Analysis and Twitter.

---

## I. INTRODUCTION

Social media is very useful in disseminating information to the public, as it is one of the most widely used media for the public to obtain knowledge. According to the We Are Social survey, by 2023, there will be 167 million people in Indonesia who are active on social media [1]. Twitter is the most widely used social media site in Indonesia, with approximately 24 million users by 2023 [2], [3], [4]. Twitter plays a significant role in shaping the narrative and public view of various emerging issues by rapidly and widely disseminating information and narratives. It also serves as a platform for challenging securitization narratives and promoting alternative perspectives [5]. Through the use of tags (#) related to a governmental event or policy, Twitter can serve as a tool for mobilizing support or protest [6]. Twitter users in Indonesia are currently discussing the 2024 presidential elections, frequently bringing up presidential candidates such as Ganjar Pranowo, Anies Baswedan, and Prabowo Subianto [7]. Many polling agencies have expressed public opinion about the 2024 presidential candidate's election. One of the many polling agencies, Charta Politika, predicted the eligibility of Ganjar Pranowo at 42.8%, Anies Baswedan at 28.1%, and Prabowo Subianto at 23.9% [8]. Public emotions about this political topic vary widely, ranging from support to hatred, and there are often hate speeches associated with presidential candidates involved in presidential elections [9]. This issue has become the main topic of discussion because the public wants to know the potential candidates to advance in the forthcoming presidential election. In a constantly changing political situation, Twitter has become the main platform for people to express their opinions and keep up with the latest developments [10].

Public posts on Twitter related to the 2024 presidential election contain a vast amount of tweet data, showcasing a variety of positive and negative views. The data collected from Twitter can provide a more comprehensive picture of voter preferences and ongoing political dynamics. This analysis can help political observers and candidates better understand the needs and expectations of the public in order to formulate an

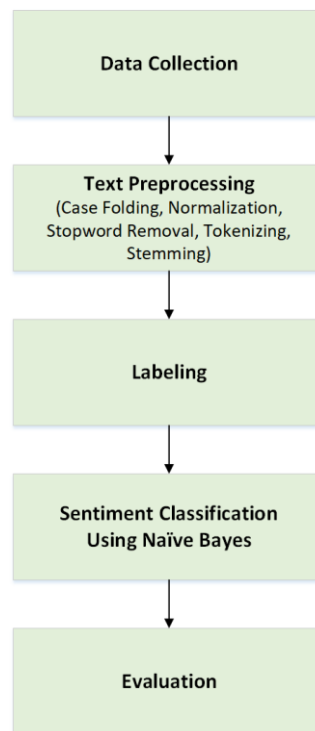
effective campaign strategy. To classify sentimental tendencies, we can process and analyze the tweet data set. We can utilize Twitter as a repository for sentimental data, allowing us to delve deeper into the information and identify patterns [11]. A society's mood can reveal the political preferences of its members [12], [13]. However, analyzing the sentiment of hundreds of thousands, even millions, of tweets takes a lot of time and effort [14]. As a result, the use of automated sentiment analysis technology can assist in processing tweet data efficiently and accurately [15], [16]. Thus, the information obtained from the sentimental analysis of tweets can provide valuable insights for decision-makers. Several previous studies have explored the use of sentiment analysis to understand public opinion and political discourse on Twitter using a variety of algorithms. Research by [17] showed the effectiveness of using the Support Vector Machine (SVM) and Word2Vec algorithms to analyze sentiment against political candidates on Twitter. However, this research is limited to the data proposition: this study mentions four data propositions but does not explain them. Although SVM is effective, it is often a "black box," and it is difficult to understand why models make specific classifications.

The research by [18] demonstrated the usefulness of Twitter data and information in generating alternative predictions and survey methods. Due to the limited time of the study, we need to conduct additional research to collect and analyze similar data during the selection period. This may indicate that Twitter has the ability to predict surveys earlier or better. The research by [19] used the K-Nearest Neighbors method for sentiment analysis, resulting in a sentiment classification accuracy rate of 86.48%. However, this study does not explain the context of tweets. For example, the frequently appearing "anias" does not reveal whether the sentiment around the candidate is positive or negative without further analysis. According to [20]'s research, Ganjar Pranowo had the most positive tweets and the most pro accounts. On the other hand, Prabowo Subianto was the subject of the most negative sentiment and controversy. However, this study does not explicitly address the collection of tweets or the use of keywords or filters. This study acknowledges the importance of language but does not clarify how the analysis reflects Indonesian cultural nuances, slang, or humor. While research mentions using Twitter data and analyzing word frequency, there is no detailed description of the methods and special tools used for data collection, sentiment analysis, and demographic analysis. Research conducted by [21] produced the Naïve Bayes model developed, showing a success rate with an accuracy of 74% for Anies Baswedan, 74% for Ganjar, and 88% for Prabowo.

However, this study ignores the sentiment of neutrality: the percentage of positive and negative feelings for each candidate does not always add up to 100%, suggesting that neutral feelings may have gone unnoticed. According to previous research, the Naive Bayes algorithm is the best for sentimental classification because it is simple to understand, faster in calculation, and requires little data training [22]. In a study using the Naïve Bayes method to analyze the sentiment towards the Telkomsel service provider on Twitter, it was found that 51 tweets had negative sentiment and 49 tweets had positive sentiment, with a category determination accuracy rate of 70.21 percent [23]. In the study that analyzed the sentiment of Twitter users towards the performance of the People's Council of Representatives using Naïve Bayes' method, it obtained a degree of precision with a score of 80% [24]. The Naïve Bayes algorithm also proved accurate in conducting an analysis of Twitter user sentiment with an accurate value of 88% [25]. The study aims to conduct a sentimental analysis of Twitter users' comments about Indonesia's 2024 presidential candidate using the Naïve Bayes algorithm. We will categorize the results of this sentimentality analysis into three classes of sentiment: positive, negative, and neutral. To evaluate the performance of Naïve Bayes, an accuracy matrix is used. We expect the research to aid in predicting public support for the presidential candidate and offer a more precise understanding of voter preferences. Thus, the results of sentimental analysis can be a strong foundation for decision-makers in designing an effective campaign strategy.

## II. METHODS

The research methodology involves several systematic steps. This methodology provides a comprehensive framework for conducting Twitter users' sentimental analysis of Indonesia's 2024 presidential candidate using the Naïve Bayes algorithm. Figure 1 shows the research measures that were used.



**Fig 1.** Research Stage

The first phase is data collection. This phase entails the collection of relevant Twitter data for the 2024 Indonesian presidential candidate. The collected data consists of tweets that mention the candidate's name, utilizing the hashtags #anies, #ganjar, and #prabowo. Data collected include tweet text, metadata such as the time of posting, the number of retweets, likes, and the identity of users. We must process the raw data from the collected tweets to prepare them for further analysis. Text preprocessing is the process that includes case folding, normalization, stopword removal, tokenizing, and stemming. Case folding will turn all letters in the text into small letters to avoid word differences caused by the use of capital letters. Normalization will remove or replace special characters and unnecessary symbols, such as over-read marks, numbers, or emoji. Stopword Removal will remove common words that do not contain important information, such as "and," "or," "with," and so on. Tokenizing will split text into words or tokens. For example, "Call of the President" becomes ["Calon," "President"]. Stemming will turn words into their basic form. For instance, stemming will transform the words "play," "play," and "play" into "play."

Following the preprocessing of the text, we label the data according to the sentiment present in the tweet. Data based on the sentiment contained in the tweet. The sentiment Machine learning algorithms, trained with previously labeled data, automatically perform the labeling have been trained with previously labelled data. The fourth phase is the classification of sentiments using the Naïve Bayes algorithm. We use this to predict sentiments from tweets using processed and labeled data. The algorithm determines the likelihood of a tweet falling into a specific sentiment category (positive, negative, or neutral) by analyzing the frequency of words in that category in the training data. This algorithm uses the Bayes rule to combine the individual probabilities of the words into the overall probability of the sentiment category. The Naïve Bayes model is trained with data that has been labeled. The model uses this data to calculate the basic probability of each word in each sentiment category. We then tested the model using new data to predict the sentiment of unlabelled tweets. The final stage is to evaluate the model's performance using accuracy metrics to ensure that it provides accurate predictions..

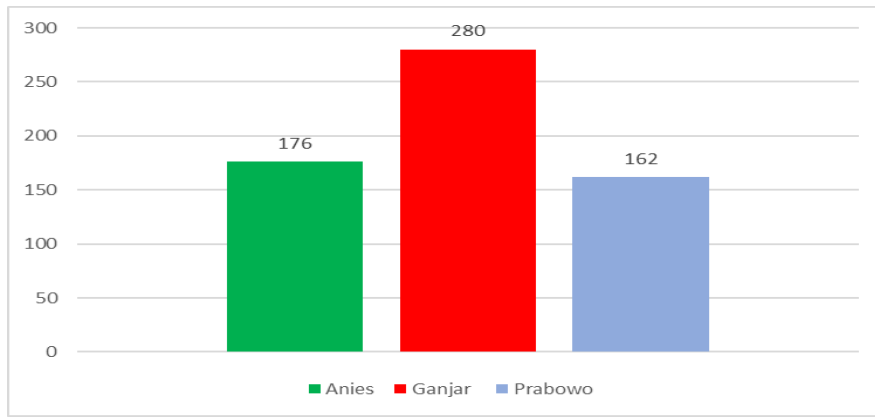
### III. RESULT AND DISCUSSION

This research applies the Naïve Bayes algorithm implemented using the Python programming language on Google Colab. The results of the research method implementation are then presented systematically in accordance with the phases of the research methodology described earlier. The data



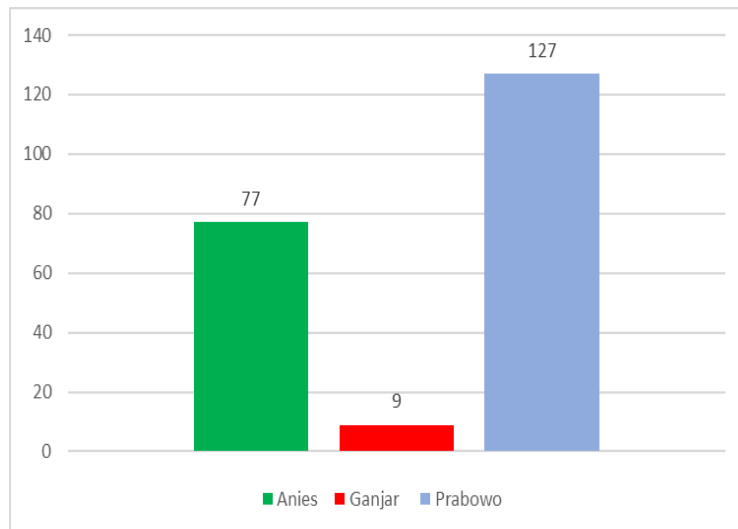






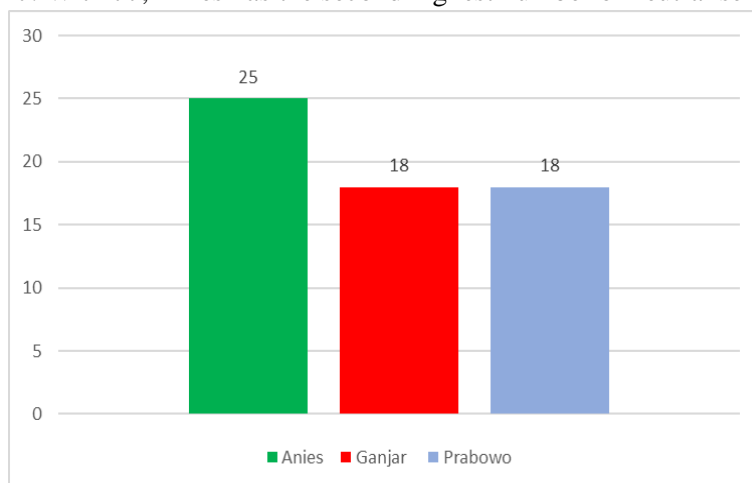
**Fig 6.** Comparison of Positive Sentiments of 2024 Presidential Candidates

Figure 6 shows a comparison of Twitter users' number of positive sentiments to the 2024 Indonesian presidential candidate. Ganjar has the highest number of positive sentiments among the three candidates, at 280. Anies has the second-highest positive sentiment with 176. Prabowo, out of the three, exhibits the least positive emotions, scoring 162.



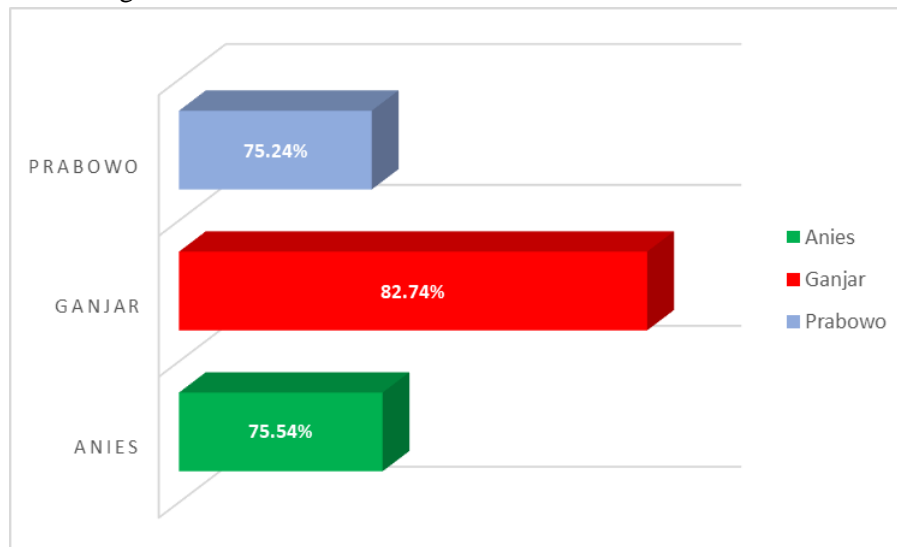
**Fig 7.** Comparison of Neutral Sentiments of 2024 Presidential Candidates

Figure 7 shows a comparison of the number of Twitter users with neutral sentiments to the 2024 Indonesian presidential candidate. Prabowo has the highest number of neutral sentiments among the three candidates, which is 127. With 77, Anies has the second highest number of neutral sentiments.



**Fig 8.** Comparison of Negative Sentiments of 2024 Presidential Candidates

Figure 8 shows a comparison of the number of Twitter users with neutral sentiments to the 2024 Indonesian presidential candidate. Anies has the highest number of negative sentiments, at 25. This shows that of the three candidates, Anies is the one who receives the most negative feedback from Twitter users. Ganjar and Prabowo share the same number of negative sentiments, totaling 18. Compared to Anies, both candidates receive less negative feedback.



**Fig 9.** Comparison of Naïve Bayes Accuracy

Figure 9 compares the accuracy of the Naïve Bayes algorithm in analyzing Twitter users' sentiment towards three Indonesian presidential candidates in 2024. Ganjar showed the highest accuracy in Twitter user sentiment analysis with a score of 82.74%. This shows that the Naïve Bayes algorithm was able to classify Twitter users' feelings towards Ganjar with a higher accurate rate than the other two candidates. Anies has an accuracy of 75.54%, and Prabowo has a precision of 75.24%. Both of these values are very close, suggesting that the Naïve Bayes algorithm has almost the same level of precision in classifying Twitter users' feelings towards Anies and Prebowo. There are quite significant differences in accuracy between Ganjar and the other two candidates. Ganjar's accuracy difference is about 7.2% higher than Anies' and 7.5% higher than Prabowo's. This research provides a strong foundation for decision-makers in designing more effective campaign strategies. By understanding public sentiment, campaign teams can adjust their messages and strategies to enhance positive sentiment and reduce negative sentiment. Knowing that Ganjar has a more easily classified sentiment, the campaign team can focus on the aspects that cause him positive or neutral sentiment and work on strategies for reducing negative sentiments. Decision makers, including political parties, campaign teams, and political analysts, can utilize the sentimental data generated by this study to make more informed decisions. The decisions made will be more targeted and aligned with public preferences and perceptions. This study could serve as the foundation for more in-depth research. For instance, we can enhance sentimental accuracy and understanding by incorporating other algorithms or employing more advanced natural language processing techniques. Additionally, one can periodically conduct a temporal analysis or trend of sentiment to observe changes in sentiment throughout the campaign period.

#### IV. CONCLUSION

The Naïve Bayes algorithm has proven effective in classifying Twitter users' sentiments towards three presidential candidates. All candidates demonstrate a fairly high accuracy of 75.54% for Anies, 82.74% for Ganjar, and 75.24% for Prabowo. High precision indicates that this algorithm is capable of effectively capturing sentiment patterns and characteristics in Twitter texts. Ganjar achieves the highest accuracy of 82.74%, indicating that this algorithm more easily classifies sentiments against Ganjar. This may be due to the existence of a clearer and more consistent pattern of sentiments towards Ganjar. We can more accurately predict public support for each presidential candidate by understanding the distribution of positive, negative, and neutral sentiments. This sentiment data can be an initial indicator of voter tendencies and potential

changes in voter preferences during campaigns. This research suggests that sentimental analysis using the Naïve Bayes algorithm can provide valuable insights into voter preferences and public support for Indonesia's 2024 presidential candidate. The results of this analysis, with their relatively high accuracy, can serve as a guide for designing more effective campaign strategies and for making more data-driven decisions. It shows the enormous potential of sentimental analysis techniques in political and campaign contexts, providing a powerful tool for decision makers to understand and respond to public sentiment more efficiently.

## REFERENCES

- [1] A. D. Riyanto, “**Hootsuite (We are Social): Indonesian Digital Report 2023**,” *andi.link*, Indonesia, Apr. 2023. [Online]. Available: <https://andi.link/hootsuite-we-are-social-indonesian-digital-report-2023/>
- [2] I. Anggorosuryo, A. R. Berto, and berta sri eko, “Menganalisis Jaringan Sosial Penggemar Blackpink Saat Konser di Jakarta,” *J. Komun. Prof.*, vol. 8, no. 1 SE-Articles, pp. 128–150, May 2024, doi: 10.25139/jkp.v8i1.7634.
- [3] Z. M. Aggriany and W. Kustiawan, “Analisis Penggunaan Media Sosial Twitter Sebagai Media Curhat oleh Kalangan Muslim Generasi Z,” *Reslaj Relig. Educ. Soc. Laa Roiba J.*, vol. 5, no. 6, pp. 3118–3133, 2023, doi: 10.47467/reslaj.v5i6.3737.
- [4] Y. Pratama, D. T. Murdiansyah, and K. M. Lhaksana, “Analisis Sentimen Kendaraan Listrik Pada Media Sosial Twitter Menggunakan Algoritma Logistic Regression dan Principal Component Analysis,” *J. Media Inform. Budidarma*, vol. 7, no. 1, pp. 529–535, 2023, doi: 10.30865/mib.v7i1.5575.
- [5] H. R. Kuncoro, D. Lupitasari, K. Hasanah, and E. Kurniawati, **MENGURAI ANCAMAN: Sekuritisasi melalui Lensa Framing dan Diskursus di Media Sosial**, 1st ed. Yogyakarta: LPPM UPNVY Press, 2023. [Online]. Available: <http://eprints.upnyk.ac.id/37540/>
- [6] R. Anbarini, S. K. A. Wibowo, N. A. Sjafrah, and A. Abdullah, “Mobilisasi Gerakan Opini Digital #OraSudiSumbangIKN pada Media Sosial Twitter,” *J. Komun.*, vol. 15, no. 1, pp. 18–36, 2023, doi: 10.24912/jk.v15i1.19187.
- [7] K. Arifin and S. I. Al-Idrus, “Klasifikasi Emosi Pengguna Twitter Terhadap Bakal Calon Presiden Pada Pemilu 2024 Menggunakan Algoritma Naïve Bayes,” *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 23, no. 1, p. 37, 2024, doi: 10.53513/jis.v23i1.9558.
- [8] A. D. Akmal, I. Permana, H. Fajri, and Yuliarti, “Opini Masyarakat Twitter terhadap Kandidat Bakal Calon Presiden Republik Indonesia Tahun 2024,” *JMIAP J. Manaj. dan Ilmu Adm. Publik*, vol. 4, no. 4, pp. 287–295, 2022, doi: doi.org/10.24036/jmiap.v4i4.160.
- [9] A. L. Efendi, A. Fadilla, A. C. Khoirunnisa, G. N. Bakry, and N. Aristi, “Analisis Jaringan Komunikasi #Pilpres2024 Pada Platform Twitter,” *WACANA J. Ilm. Ilmu Komun.*, vol. 22, no. 2, pp. 219–232, 2023, doi: 10.32509/wacana.v22i2.2976.
- [10] M. A. Hayat, S. Jayadiningrat, G. Wibisono, and M. I. Iyansyah, “Peran Media Sosial Dalam Komunikasi Politik,” *J. Indones. Sos. Teknol.*, vol. 2, no. 1, pp. 104–114, 2021, doi: 10.36418/jist.v2i1.61.
- [11] R. Amelia, D. Darmansah, N. S. Prastiwi, and M. E. Purbaya, “Impementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Masyarakat Indonesia Mengenai Drama Korea Pada Twitter,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 2, p. 338, 2022, doi: 10.30865/jurikom.v9i2.3895.
- [12] A. A. Arifiyanti, E. D. Wahyuni, and A. Kurniawan, “Emotion Mining of Indonesia Presidential Political Campaign 2019 using Twitter Data,” *J. Phys. Conf. Ser.*, vol. 1569, no. 2, p. 22035, 2020, doi: 10.1088/1742-6596/1569/2/022035.
- [13] F. D. N. Putra, Pranowo, and B. Setyohadi, “Sentiment analysis of Indonesian presidential election 2019 on the twitter with lexicon-based and support vector machine (SVM),” *AIP Conf. Proc.*, vol. 2217, no. 1, p. 30136, Apr. 2020, doi: 10.1063/5.0000631.
- [14] R. Vindua and A. U. Zailani, “Analisis Sentimen Pemilu Indonesia Tahun 2024 Dari Media Sosial Twitter Menggunakan Python,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 10, no. 2, p. 479, 2023, doi: 10.30865/jurikom.v10i2.5945.
- [15] A. Kaharudin, A. A. Supriyadi, Muhlis, H. Baitika, and M. Derryanur, “Analisis Sentimen pada Media Sosial dengan Teknik Kecerdasan Buatan Naïve Bayes: Kajian Literatur Review,” *OKTAL J. Ilmu Komput. dan Sci.*, vol. 2, no. 6, pp. 1642–1649, 2023.
- [16] A. Faisal, Y. Alkhalifi, A. Rifai, and W. Gata, “Analisis Sentimen Dewan Perwakilan Rakyat dengan Algoritma



- Klasifikasi berbasis Particle Swarm Optimization,” *JOINTECS (Journal Inf. Technol. Comput. Sci.*, vol. 5, no. 2, p. 61, 2020, doi: 10.31328/jointecs.v5i2.1362.
- [17] L. Damayanti and K. M. Lhaksana, “Sentiment Analysis of the 2024 Indonesia Presidential Election on Twitter,” *Sinkron*, vol. 8, no. 2, pp. 938–946, 2024, doi: 10.33395/sinkron.v8i2.13379.
- [18] T. Baharuddin, Z. Qodir, H. Jubba, and A. Nurmandi, “Prediction of Indonesian presidential candidates in 2024 using sentiment analysis and text search on Twitter,” *Int. J. Commun. Soc.*, vol. 4, no. 2, pp. 204–213, 2022, doi: 10.31763/ijcs.v4i2.512.
- [19] K. Diantoro, A. Soderi, A. Rohman, and A. T. Sitorus, “Sentiment Analysis of Public Opinion on the 2024 Presidential Election in Indonesia Using Twitter Data with the K-NN Method,” *Digit. J. Comput. Sci. Appl.*, vol. 1, no. 1, pp. 1–10, 2023, doi: 10.61978/digitus.v1i1.27.
- [20] N. U. Rahmanulloh and I. Santoso, “Delineation of The Early 2024 Election Map: Sentiment Analysis Approach to Twitter Data,” *JOIN (Jurnal Online Inform.*, vol. 7, no. 2, pp. 226–235, 2022, doi: 10.15575/join.v7i2.925.
- [21] R. Hakiki, A. Pambudi, and Asriyanik, “Classification of Public Sentiment Toward 2024 Presidential Candidates on Social Media Platform X Using Naïve Bayes Algorithm,” *J. Artif. Intell. Eng. Appl.*, vol. 3, no. 2, pp. 551–556, 2024, doi: 10.59934/jaiea.v3i2.422.
- [22] N. A. Salsabila, “Analisis Sentimen pada Media Sosial Twitter terhadap Tokoh Gus Dur menggunakan Metode Naive Bayes dan Support Vector Machine (SVM),” UIN Syarif Hidayatullah, 2022.
- [23] N. Haqqizar and T. N. Larasyanti, “Analisis Sentimen Terhadap Layanan Provider Telekomunikasi Telkomsel Di Twitter Dengan Metode Naïve Bayes,” in *Prosiding TAU SNAR-TEK 2019 Seminar Nasional Rekayasa dan Teknologi*, 2019, pp. 1–15.
- [24] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, “Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier,” *J. Inform. dan Tek. Elektro Terap.*, vol. 10, no. 1, pp. 34–40, 2022, doi: 10.23960/jitet.v10i1.2262.
- [25] H. M. Saragih, “Analisis Sentimen Pengguna Twitter Terhadap Layanan Pajak Kendaraan Bermotor Menggunakan Algoritme Naive Bayes Classifier,” Universitas Lampung, 2021.