



Research Article

Open Access (CC-BY-SA)

Analysis of public opinion on COVID-19 vaccine through social media using Naïve Bayes theory algorithm

Aishiyah Saputri Laswi^{a,1,*}, Munir Yusuf^{a,2}, Ulvah^{b,3}, Bungawati^{a,4}

^a Institut Agama Islam Negeri Palopo, Jl. Agatis, Palopo and 91914, Indonesia

^b Universitas Cokroaminoto Palopo, Jl. Lamatacelling No. 19, Palopo and 91911, Indonesia

¹ aishiyahsaputrilaswi@iainpalopo.ac.id; ² munir_yusuf@iainpalopo.ac.id; ³ ulvahinformatik@gmail.com; ⁴ bungawati@iainpalopo.ac.id

* Corresponding author

Article history: Received June 04, 2022; Revised June 07, 2022; Accepted July 02, 2022; Available online August 30, 2022

Abstract

This study aims to analyze various public opinion on the Covid-19 vaccines that appear on social media pages, especially on Facebook and Twitter via # (hashtag). The death rate caused by COVID-19 was so high which reached 144,227 people until 2022. The Indonesian government required vaccines for the community starting from children aged 6 years as an effort to prevent the spread of the Covid-19 virus. Unfortunately, the implementation of complete vaccines in Indonesia has only reached 51.3% of the mandatory vaccine population, which is 140 million out of 339 million people. The non-achievement of the target set by the government causes the need to conduct a sentiment analysis on vaccines in Indonesia through social media. Based on the sample data, from 1000 words obtained from 320 opinions there are positive and negative opinions. This data is then analyzed and processed to find out how many positive and negative responses occurred. The data was then processed into several stages to test the level of truth through training data and test data. The results of the data processing were tested using the Naïve Bayes algorithm which resulted in an accuracy value with a precession of 77.08% taken from 90 samples test data, recall with a percentage of 97.87% based on positive data which was predicted to be true with a positive opinion status from 47 samples of test data and 1 positive data status which is still predicted to be negative. Furthermore, the specific percentage value obtained was 65.30% of the 132 test data that are predicted to be true negative.

Keywords: Analysis; Opinion; Social Media; Naïve Bayes

Introduction

Covid-19 is a disease that attacked the world at the end of 2019, firstly appeared in Wuhan, a city located in China [1]–[3],[4]. The disease caused by this deadly virus is called the Corona Virus. This new variant of the virus attacks the human respiratory system through the lungs. The disease that first appeared in Wuhan is increasing along with the growing number of deaths in the city. Based on data from the WHO, Covid-19 cases continue to increase with higher positive cases in various parts of the world. Consequently, Covid-19 cases have received special attention from the World Health Organization (WHO), which is a world health organization established by the United Nations. March 2020 was declared a pandemic period caused by Corona Virus Disease 19 (Covid-19)[3], [5]–[7].

In Indonesia, a positive case of Covid-19 was officially announced on March 2, 2020 based on the confirmation results from two people in the first case who tested positive for Covid-19.

On March 13, other new cases of 35 people who were confirmed positive for Covid-19 were reported. Since then, the number of positive cases has continued to increase every day, with the total daily cases ranging from 200 to 400 new cases. This led the government to strictly implement a National Lockdown in Indonesia [7], [8].

The determination of Covid-19 as a global health emergency led the Indonesian government to issue Presidential Decree of the Republic of Indonesia Number 99 of 2020 concerning the Procurement of Vaccinations for Combating the Corona Virus Diseases (Covid-19) Pandemic which was officially announced by the President of Indonesia on October 5, 2020 [9],[10]. Ministries/Institutions and their officials as well as civil servants were the

first targets to receive free vaccines. This Presidential Decree was immediately followed up by all elements involved. On December 6, 2020, Indonesia brought in the Cornavac Dose Vaccine produced by Sinovac Biotech from Beijing and directly registered with PT. Bio Farma Bandung[10].

With the national vaccine decision that has been officially announced by the president, the plan for vaccine activities begins by involving several relevant government elements and the wider society by considering the stages and procedures in administering vaccines, starting from the feasibility of the vaccine to the risks after the vaccine. This vaccine has gone through several stages of clinical trials by the Food and Drug Supervisory Authority (BPOM) and the Indonesian Ulama Council (MUI). After being declared to have obtained quality feasibility and halal testing from MUI, on January 13, 2021 the first vaccination was carried out. [1], [2],[10].

The initial vaccine carried out caused various responses and opinions from the Indonesian people which were expressed through social media. [3], [6], [11]–[13]. Opinions expressed through social media are growing faster without limits of distance and time so that it could create a commotion in the community which caused a lot of direct rejection from the wider community of vaccines. This refusal made the government have to take effort to ensure the society that the COVID-19 vaccine is not dangerous, instead, it is one way to protect society from the COVID-19 virus as well as increase the immune system so that they are easily attacked by the Covid-19 virus. It is important to know the actual vaccine facts in order to avoid mistaken opinions which are strongly influenced by a person's feelings, perspectives, attitudes, experiences and understandings. Therefore, this study collects data that appears on social media pages via #(hashtag) [14],[15].

This study discusses public opinion on the Covid-19 Vaccine circulating on social media using the Naïve Bayes Method. This method is a classification technique used to predict the frequency of occurrence of a word in a sentence. In the context of this study, the Naïve Bayes method is used to analyze every public opinion that appears on social media pages, whether it is positive, negative, or neutral, sourced from Facebook and Twitter through hashtags (#) which are then classified into positive and negative opinions to provide an illustration to the public that the Covid-19 vaccine is actually not dangerous and many support this vaccine program.[16],[17][18].

Methods

The sample of this research is public opinion that appears on social media pages through #(hashtag) regarding the procurement of the covid-19 vaccine held by the government. There were many rejections from the community at the beginning because of negative issues that had already spread widely on Social Media pages that made a public scene. Several stages conducted in this study as follows:

A. Data Collection

The sampling process is carried out by utilizing the Application Interface (API) provided by several social media such as Facebook and Twitter. This API application functions to find updated status on social media walls that have been uploaded by users in the last 7 days. From this process, words related to the covid -19 vaccine that most often appear on social media pages using the hashtag(#) could be obtained [19].

The words that appear are then processed to be used as test data. The collected data used as a sample for training are neutral, positive and negative values based on the intensity of the word appearing through the hashtag (#) on the social media page shown in **Table 1**. The word used as a sample for the test data will then be processed to determine the TDF and IDF which serves to determine the distance that appears in each word contained in the document by combining two concepts for weight calculation, namely the frequency of occurrence of a word in a document and the inverse which contains the word. The method used is the Term Frequency – Inverse Document Frequency method or commonly called TFIDF.

Table 1. Sample of Test Data

Test Data Sample			
<i>True Positive</i>	<i>False Positive</i>	<i>True Negative</i>	<i>False Negative</i>
D1	D2	D3	D4
Nafsu	mari,	Bahaya	Menolak
Tidur	Sukseskan	Haram	jangan
Kebal	Ayo	Palsu	Bodoh
Vaksin	Dukung	Cina	astra Zeneca
Sehat	mulai dari	akal pemerintah	tidak
BB naik	lawan	Kecuali	Pidana
Makan	Program	Konspirasi	Ancaman

Test Data Sample			
True Positive	False Positive	True Negative	False Negative
	Pemerintah	takut	Pemaksaan
	Halal	Mati	
	Covid	Sakit	

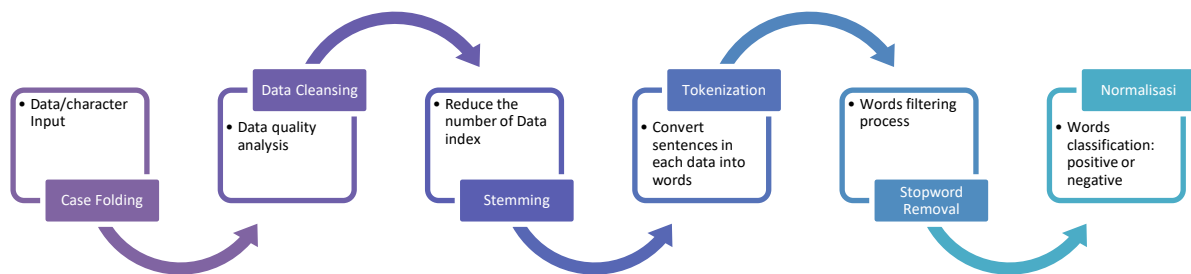
B. Data Processing

The data that has been obtained were then be analyzed. Before that, the data were displayed again. Any inappropriate data had to be removed or deleted to make it easier for the data analysis stage so that the processed data can work efficiently. Several structures available on social media (Facebook and Twitter) such as text, date, time, file, ID, username, image and video as well as hashtags have their respective levels of relevance during the analysis process. Some of the data taken from the opinion does not have any relevant power so that it is eliminated during the analysis process.

C. Data Analysis

Data analysis was carried out by grouping public opinion on the covid-19 vaccine appeared on Facebook and Twitter through several stages, consisting of case folding, data cleansing, stemming, tokenization, stopword removal and normalization. [20],[21]. Based on **Figure 1** is Stages of Data Analysis.

Figure 1. Stages of Data Analysis



D. Word weighting

Word weighting is to assign a distance value that appears on each word in a document. Giving value to each word is done by the method of Term Frequency – Inverse Document Frequency (TFIDF), [8], [11], which is a statistical numbering method to analyze a word using the following **Equation 1**:

$$TF - IDF = TF \times IDF = TF_t, x = \frac{|D|}{DF_t} \quad (1)$$

Note:

TF_t, d : The total frequency of occurrence of the word t in a document d

DF_t : Total documents that store the word t

D : Total of all available documents

TF : The number of words searched for in a document

IDF : Inverse Document Frequency

Table 2 is a TF-IDF table by taking samples from the test data in **Table 1**.

Table 2. Calculation of TF-IDF

TERM	DOCUMENTS				DF	IDF	TF-IDF			
	D1	D2	D3	D4			D1	D2	D3	D4
Nafsu	1	0	0	0	1	0,602	0,602	0	0	0
Tidur	1	0	0	0	1	0,602	0,602	0	0	0
Kebal	1	0	0	0	1	0,602	0,602	0	0	0
Vaksin	1	1	1	1	4	0	0	0	0	0
Sehat	1	0	0	0	1	0,602	0	0	0	0
BB naik	1	0	0	0	1	0,602	0	0	0	0
Makan	1	0	0	0	1	0,602	0	0	0	0
mari,	0	1	0	0	1	0,602	0	0,602	0	0

TERM	DOCUMENTS				DF	IDF	TF-IDF			
	D1	D2	D3	D4			D1	D2	D3	D4
Sukseskan	0	1	0	0	1	0,602	0	0,602	0	0
Ayo	0	1	0	0	1	0,602	0	0,602	0	0
Dukung	0	1	0	0	1	0,602	0	0,602	0	0
mulai dari	0	1	0	0	1	0,602	0	0,602	0	0
Lawan	0	1	0	0	1	0,602	0	0,602	0	0
Program	0	1	0	0	1	0,602	0	0,602	0	0
Pemerintah	0	1	0	0	1	0,602	0	0,602	0	0
Halal	0	1	0	0	1	0,602	0	0,602	0	0
Covid	0	1	1	1	3	0,200	0	0,200	0	0
Bahaya	0	0	1	1	2	0,301	0	0	0,301	0
Haram	0	0	1	1	2	0,301	0	0	0,301	0
Palsu	0	0	1	1	2	0,301	0	0	0,301	0
Cina	0	0	1	1	2	0,301	0	0	0,301	0
akal pemerintah	0	0	1	0	1	0,602	0	0	0,602	0
Kecuali	0	0	1	0	1	0,602	0	0	0,602	0
Konspirasi	0	0	1	0	1	0,602	0	0	0,602	0
Takut	0	0	1	0	1	0,602	0	0	0,602	0
Mati	0	0	1	0	1	0,602	0	0	0,602	0
Sakit	0	0	1	0	1	0,602	0	0	0,602	0
Menolak	0	0	1	1	2	0,301	0	0	0	0,301
jangan	0	0	0	1	1	0,602	0	0	0	0,602
Bodoh	0	0	0	1	1	0,602	0	0	0	0,602
astra Zeneca	0	0	0	1	1	0,602	0	0	0	0,602
Tidak	0	0	0	1	1	0,602	0	0	0	0,602
Pidana	0	0	0	1	1	0,602	0	0	0	0,602
Ancaman	0	0	0	1	1	0,602	0	0	0	0,602
pemaksaan	0	0	0	1	1	0,602	0	0	0	0,602

Note:

Term : Document term

D1 : Document 1

D2 : Document 2

D3 : Document 3

D4 : Document 4

After calculating the data on the training data and test data, then the data will be divided randomly and balanced based on 1000 words out of a total of 320 data based on opinions that appear on social media pages via the hashtag (#)

Table 3. Data Sharing

Training Data		Test Data	Total
Label PF	132	50	182
Label NF	98	40	138
Total	230	90	320

Table 3 is a separation of Training Data and Test Data based on False Positive Labels and False Negative Labels which can be seen through the graph in **Figure 2** below:

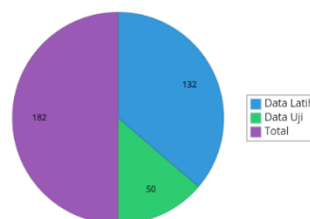


Figure 2. Distribution of Test Data and Training Data Graph

E. Evaluation

Based on testing the results of the data that have been predicted previously, the following configuration table shows the percentage value of 2 class categories that are right on target using correct answers:

Table 4. Data Sharing

Prediction	Positive	Negative
Positive	True Positive (TP)	False Negative (FP)
Negative	False Negative (FN)	True Negative (TN)

Note:

True Positive (TP), opinion with actual positive and predicted positive

True Negative (TN), opinion with actual negative and predicted negative

False Positive (FP), opinion with actual negative but predicted positive

False Negative (FN), opinion with actual positive but predicted negative

Table 5. Prediction Evaluation Results

Prediction		
	Positive	Negative
Positive	46 TP	21 FP
Negative	1 FN	28TN

Table 5 is the result of the evaluation of the predictions that have been given in **Table 4**. The data was then used to determine the precession value to answer the positive correct prediction of the overall positive data. This is to determine the percentage of the correct answer from the overall opinion, as well as precession, recall and specifics. Precession value is calculated by the following **Equation 1**:

$$Precession = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

The recall value is determined using the following **Equation 2**:

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

Value of specific is determined using the following **Equation 3**:

$$Specific = \frac{TN}{TN + FP} \quad (3)$$

Table 6. Evaluation Calculation Results

Measurement	Value
Accuracy	77,08%
Recall	97,87%
Specific	65,30%

Table 6 shows the results of the evaluation calculations using the Naïve Bayes model. Based on these results, the system can classify the Covid-19 vaccine opinion analysis by determining the Accuracy, Recall and Specific values as depicted in **Figure 3**. The test results using Naïve Baye produced an accuracy value of 77.08% which is the ratio of true positive and negative predictions of a total of 90 test data. The recall value is 97.87% of the positive data that is correctly predicted positive with a total of 47 data. While the specific value which is the ratio of negative predictions by comparing the overall negative data is 65.30%

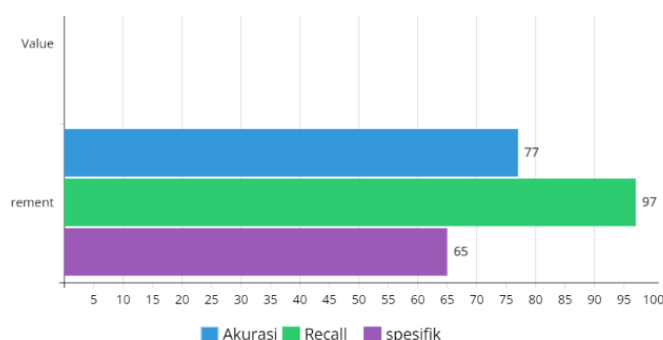


Figure 3. Evaluation Calculation Results

Results and Discussion

Based on research on public opinion on the covid-19 vaccine that has been studied using the Naïve Bayes algorithm, there are several negative opinions that have emerged in the community that spread so quickly through social media pages, both Facebook and twitter. Using the search keywords #CoronaVaksin, #VaksinCovid-19 and Vaccinesiovac, we found Positive Opinions and Negative Opinions on Social Media pages from 1000 words that had been taken from a total of 320 opinions sampled. The data is then processed in several stages to test the level of truth through training data and test data.

On the other hand, based on the Naïve Bayes test, the accuracy value with precession is 77.08%, obtained from the total test data of 90 samples. The recall percentage value is 97.87% based on positive data which is predicted to be correct with a positive opinion status with a total of 47 samples of test data with 1 positive data status which is still predicted to be negative. The larger the False Positive (FP), the smaller the precession accuracy value and the smaller the False Negative (FN) value, the greater the Recall value. The Precession Accuracy value is lower than the Recall value because the Precession Accuracy value has a significant effect on the Recall value, but the interpretation of the Precession Accuracy and Recall values is categorized as valid.

Furthermore, the percentage value of the Specific obtained is 65.30% of the 132 test data that are predicted to be true negative. Words that often appear in positive opinions about the COVID-19 vaccine are 'vaksin', 'covid', 'vaksinasi', 'Indonesia', 'sehat', 'terima', 'aman', etc. The word that most often appears in positive sentiments about the COVID-19 vaccine is 'vaccine' as many as 189 words. While the words that most often appear in negative opinions about the COVID-19 vaccine are 'vaksin', 'covid', 'tolak', 'sinovac', 'hoaks', 'suntik', etc. The word that most often appears in negative opinions about the COVID-19 vaccine is 'vaccine' as many as 147 words as shown in figure 4 below.

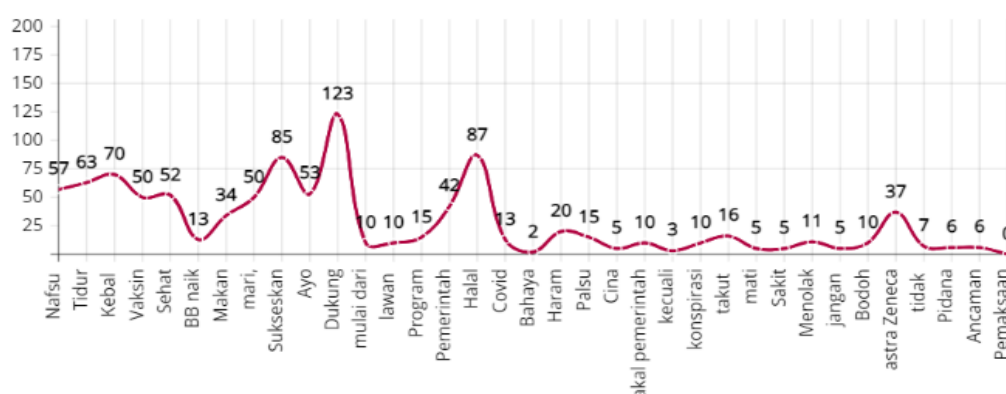


Figure 4. Graph of Test Data Sample

'Vaccine' is the word that appears most frequently in the 2 categories below #nafsu(57), #makan(34), #tidur(63), #kebal(50), #sehat(52), #BBNaik(13), #mari(50), #dukung(123), #mulaidari(10), #lawan(10), #program(10), #pemerintah(42), #halal(87), #covid(13). These words that often appear are positive opinions from some comments via #(Hashtag). In addition to positive opinions, there are also many people who do not receive vaccine services by raising various negative opinions on social media pages that also use #(hashtag) such as

#bahaya(2), #haram(20), #palsu(15), #cina(5), #akalpemerintah(10), #konspirasi(10), #takut(16), #mati(5), #sakit(5), #menolak(11), #jangan(5), #bodoh(10), #AstraZeneca(37), #tidak(7), #pidana(6), #ancaman(6). This data is the result of a search using # (hashtag) which is then used as a sample for training data and test data to produce the TD-IDF value which is the word weighting before the evaluation stage. The IDF value has passed 2 stages, namely training data and test data. The lowest and highest IDF value are 0 and 0.602 respectively. Based on these results, the following data are used as samples to analyze opinions that appear on the social media page.

Table 6. Opinion

Opinion	Label
<i>Ayo Vaksin Insha Allah aman!</i> (Come on, vaccine insya Allah safe!)	Positif
<i>Mari bersama dukung Vaksin Covid-19</i> (Let's support the Covid-19 Vaccine together)	Positif
<i>Akal Pemerintah buat membodohi rakyat</i> (The government's sense to fool the people)	Negatif
<i>Vaksin itu halal</i> (Vaccine is halal)	Positif
<i>Semoga dengan adanya vaksin sebagai salah satu cara buat mengurangi penyebaran covid 19</i> (Hopefully, vaccine is one way to decrease the number of the spread of covid-19)	Positif
<i>Tolak Vaksin</i> (Reject vaccine)	Negatif
<i>Vaksin bagian dari konspirasi pemerintah</i> (Vaccine is part of government conspiracy)	Negatif
<i>Salah satu cara pemerintah buat korupsi dengan pengadaan vaksin</i> (One of the ways of government committing corruption is by procuring vaccine)	Negatif
<i>Kami saja sudah di vaksin. Yuk, kamu juga segera di vaksin ya!</i> (We've just been vaccinated. Come on, you too get vaccinated right away!)	Positif
<i>Vaksin bisa bikin orang yang sehat malah tambah sakit</i> (Vaccines can make healthy people sicker)	Negatif
<i>Demam tinggi setelah vaksin</i> (High fever after vaccine)	Negatif
<i>Jangan mau dibodohi pemerintah dengan program vaksin</i> (Don't be fooled by the government with a vaccine program)	Negatif
<i>Vaksin mengandung minyak bumi</i> (Vaccine contains p*rk)	Negatif
<i>Program vaksin ini merupakan pemaksaan bagi masyarakat</i> (Vaccination program is a compulsion for society)	Negatif
<i>Habis vaksin bawaanya pengen makan terus</i> (After getting vaccinated, I naturally want to eat)	Positif
<i>BB Naik setelah Vaksin</i> (Gaining weight after vaccination)	Positif
<i>Sehat setelah Vaksin</i> (Being healthy after vaccination)	Positif
<i>Vaksin aman kok!</i> (Vaccine is safe!)	Positif
<i>Vaksin dapat melindungi kekebalan tubuh dari virus</i> (Vaccines can protect the immune system from viruses)	Positif
<i>Vaksin program pemerintah untuk meredem penyebaran covid</i> (The government's vaccine program is to reduce the spread of covid)	Positif

Table 6 is the result of an analysis of public opinion about the covid-19 vaccine. This study shows how important the role of social media is in today's social life. Therefore, the government needs to anticipate this through education on the introduction of vaccines before holding a covid-19 vaccine so that there are no more negative things or opinions that appear on social media pages that may cause commotion in the society as previously happened.

Conclusion

This study examines public opinion on the Covid-19 vaccine circulating through Social Media pages, which cannot be limited. The results showed that there were many positive opinions that appeared on social media pages through # (Hashtag). Even though at the beginning of the emergence of this vaccine, there were many rejections from various circles of society and even elements of the government so that it had caused a commotion on social media. Of the 1000 words taken from 320 opinions that were used as the test sample, there are opinions that are Negative True, Positive True, Negative False, Positive False.

The Naïve Bayes method is used to analyze public opinion on social media pages to measure negative opinions that are true, negative opinions are false, positive opinions are false and positive opinions are true by using the TF-IDF method. This stage produces test data and training data which is the source of accuracy, recall and specific values to produce positive and negative opinion classifications as shown in **Figure 3**. The percentage of accuracy, specific and recall values calculated is 77.08%, 97.87% and 65.30% respectively. The recall value is the correctness of predicting the negative compared to the overall negative data. By using the Naïve Bayes Theory Algorithm, the accuracy of opinions can have a reliable level of validity based on the results shown in **Table 3**. A previous study,

Twitter Social Media Text Classification Using Support Vector Mechine, was lack in terms of inability to read foreign words contained in a text. sentence for each opinion that appears [22].

The results of this study can be used by the government as evaluation considering various opinions that arise in the society. By doing so, the government can, hopefully, increase public confidence in the Covid-19 vaccine. It is suggested for the government to pay attention to the quality of the types of vaccines used and to be more careful in conducting health screenings for people who will carry out vaccines so as not to cause negative opinions from public.

References

- [1] A. Harun and D. P. Ananda, "Analysis of public opinion sentiment about COVID-19 vaccination in Indonesia using Naïve Bayes and Decision Tree Analisa Sentimen opini publik tentang vaksinasi COVID-19 di Indonesia menggunakan Naïve Bayes dan Decission Tree," *Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. April, pp. 58–63, 2021.
- [2] A. Baita, Y. Pristyanto, N. Cahyono, P. Covid-, K. N. N. Akurasi, and K. Kunci, "analisis sentimen mengenai vaksin sinovac menggunakan algoritma Support Vector Machine (SVM) dan K-Nearst Neighbor (KNN) abstraksi keywords :," vol. 4, no. 2, pp. 42–46, 2021.
- [3] S. Lestari and S. Saepudin, "Analisis sentimen vaksin sinovac pada twitter menggunakan algoritma Naive Bayes," *SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform.*, pp. 163–170, 2021.
- [4] S. S. Aljameel *et al.*, "A sentiment analysis approach to predict an individual's awareness of the precautionary procedures to prevent covid-19 outbreaks in Saudi Arabia," *Int. J. Environ. Res. Public Health*, vol. 18, no. 1, pp. 1–12, 2021, doi: 10.3390/ijerph18010218.
- [5] B. Laurensz and Eko Sedyono, "Analisis sentimen masyarakat terhadap Tindakan vaksinasi dalam upaya mengatasi pandemi COVID-19," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 2, pp. 118–123, 2021, doi: 10.22146/jnteti.v10i2.1421.
- [6] D. Sandi, E. Utami, and A. Fatkhurohman, "Klasifikasi opini dengan menggunakan algoritma K- Nearest Neighbor pada berita vaksinasi di twitter," vol. 16, 2022.
- [7] F. Fitriana, E. Utami, and H. Al Fatta, "Analisis sentimen opini terhadap vaksin Covid - 19 pada media sosial twitter menggunakan Support Vector Machine dan Naive Bayes," *J. Komtika (Komputasi dan Inform.*, vol. 5, no. 1, pp. 19–25, 2021, doi: 10.31603/komtika.v5i1.5185.
- [8] Merinda Lestandy, Abdurrahim Abdurrahim, and Lailis Syafa'ah, "Analisis sentimen tweet vaksin COVID-19 menggunakan Recurrent Neural Network dan Naïve Bayes," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 802–808, 2021, doi: 10.29207/resti.v5i4.3308.
- [9] President of the Republic of Indonesia, "Peraturan Presiden No. 99 Tahun 2020 tentang pengadaan vaksin dan pelaksanaan vaksinasi dalam rangka penanggulangan pandemi Corona Virus Disease 2019," *Pres. Regul.*, vol. 2019, no. 039471, pp. 1–13, 2020, [Online]. Available: <https://peraturan.bpk.go.id/Home/Details/147944/perpres-no-99-tahun-2020>.
- [10] Government of Indonesia, "Presidential Decree Number 99," vol. 2019, no. 1, pp. 1–8, 2020.
- [11] W. Yulita *et al.*, "Analisis sentimen terhadap opini masyarakat tentang vaksin COVID-19 menggunakan algoritma Naïve Bayes Classifier," *J. Data Min. dan Sist. Inf.*, vol. 2, no. 2, pp. 1–9, 2021, [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/JDMSI/article/view/1344>.
- [12] P. Arsi, L. N. Hidayati, and A. Nurhakim, "Komparasi model klasifikasi sentimen issue vaksin COVID-19 berbasis platform instagram," vol. 6, pp. 459–466, 2022, doi: 10.30865/mib.v6i1.3509.
- [13] A. Putra, D. Haeirudin, and H. Khairunnisa, "Analisis sentimen masyarakat terhadap kebijakan PPKM pada Media Sosial Twitter Menggunakan Algoritma Svm," no. November, 2021.
- [14] P. D. Bangsa and I. Hermawan, "Jurnal Teknologi Terpadu," *J. Teknol. Terpadu*, vol. 7, no. 1, pp. 15–22, 2021.
- [15] B. Bahekar, P. Gautam, and S. Sharma, "Literature review on sentiment analysis and opinion classifications on the impact of COVID19 outbreak," vol. 9, no. 1, pp. 94–99, 2022.
- [16] S. Lorena., "Teknik data mining menggunakan metode Bayes Classifier untuk optimalisasi pencarian aplikasi perpustakaan," *J. Tek. Komput.*, vol. 4, no. 2, pp. 17–20, 2016.

-
- [17] R. E. Putri, Suparti, and R. Rahmawati, "Perbandingan metode klasifikasi Naïve Bayes dan K-Nearest Neighbor pada analisis data status kerja di kabupaten demak tahun 2012," *J. Gaussian*, vol. 3, no. 4, pp. 831–838, 2014.
- [18] N. Imtiaz Khan, T. Mahmud, and M. Nazrul Islam, "COVID-19 and black fungus: Analysis of the public perceptions through machine learning," *Eng. Reports*, no. September, pp. 1–10, 2021, doi: 10.1002/eng2.12475.
- [19] E. S. Negara, R. Andryani, and P. H. Saksono, "Analisis data twitter: ekstraksi dan analisis data geospasial," *J. INKOM*, vol. 10, no. 1, p. 27, 2016, doi: 10.14203/j.inkom.433.
- [20] D. D. Palmer, "Text Pre-processing," *Handb. Nat. Lang. Process. Second Ed.*, 2010.
- [21] D. Rustiana and N. Rahayu, "Analisis sentimen pasar otomotif mobil: tweet twitter menggunakan Naïve Bayes," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 8, no. 1, pp. 113–120, 2017, doi: 10.24176/simet.v8i1.841.
- [22] L. Mutawalli, M. T. A. Zaen, and W. Bagye, "Klasifikasi teks sosial media twitter menggunakan Support Vector Machine (Studi Kasus Penusukan Wiranto)," *J. Inform. dan Rekayasa Elektron.*, vol. 2, no. 2, p. 43, 2019, doi: 10.36595/jire.v2i2.117.