Spirit Publik Volume 17, Nomor 1, 2022
Halaman 76-87
P-ISSN. 1907-0489 E-ISSN 2580-3875

Sentiment Analysis of Covid-19 Vaccination Policy in Indonesia

Analisis Sentimen Terhadap Kebijakan Vaksinasi Covid-19 di Indonesia

Ahmad Zaky, Retno Kusumastuti
Fakultas Ilmu Administrasi, Universitas Indonesia
Email: ahmad.zaky01@ui.ac.id, r.kusumastuti@ui.ac.id

Abstrak
Pada bulan Desember tahun 2020 pandemi covid-19 sudah berlangsung selama 10 bulan di Indonesia dan belum terlihat jalan keluar dari pandemi tersebut. Secercah harapan datang dari Presiden Joko Widodo pada bulan Agustus 2020 ketika menyampaikan bahwa program vaksinasi akan dimulai pada bulan Januari 2021 dan vaksin sudah tiba di Indonesia pada bulan Desember 2020. Beragam tanggapan muncul di masyarakat terhadap kebijakan vaksinasi tersebut, terutama tanggapan yang disampaikan di media sosial. Kepercayaan publik terhadap suatu kebijakan merupakan salah satu faktor yang mendorong keberhasilan suatu kebijakan. Artikel ini melihat bagaimana sentimen masyarakat terhadap kebijakan vaksinasi covid-19 dengan melakukan text mining terhadap komentar warganet di media sosial youtube dan menganalisis sentimen yang diberikan warganet dan menganalisis topik-topik apa saja yang dibicarakan seputar kebijakan vaksinasi covid-19. Hasil penelitian menunjukkan bahwa mayoritas warganet memiliki sentimen negatif terhadap kebijakan vaksinasi covid-19 dengan kekhawatiran antara lain seputar masalah keamanan vaksin dan potensi korupsi dari kebijakan tersebut. Untuk meningkatkan kepercayaan publik terhadap kebijakan vaksinasi covid-19, pemerintah harus menjawab kekhawatiran masyarakat tersebut. Kekhawatiran terhadap keamanan vaksin dapat diatasi dengan memberikan contoh atau tauladan dari pejabat tinggi pemerintah dengan bersedia divaksinasi terlebih dahulu dan disiarkan secara langsung sebagaimana dilakukan oleh presiden dan wakil presiden terpilih Amerika Serikat Joe Biden dan Kamala Haris. Sedangkan untuk masalah kekhawatiran terhadap potensi korupsi dapat diatasi dengan melakukan tindakan-tindakan yang lebih transaparan, akuntabel dan responsif dalam rangka menjalankan kebijakan vaksinasi.

Kata Kunci : Analisis Sentimen; Kebijakan Publik; Kepercayaan Publik; Vaksinasi Covid-19

Abstract
In December 2020, the COVID-19 pandemic had lasted for 10 months in Indonesia and there was no way out of the pandemic. A glimmer of hope came from President Joko Widodo in August 2020 when he said that the vaccination program would start in January 2021 and the vaccine had arrived in Indonesia in December 2020. Various responses have emerged in the community regarding the vaccination policy, especially on social media. Public trust in a policy is one of the factors driving a successful policy. This article shows how the public's sentiment is towards the covid-19 vaccination policy by conducting text mining on netizen comments on YouTube social media, analyzes the sentiment from citizens, and analyzes the topics discussed about the covid-19 vaccination policy. The results showed that the majority of netizens had negative sentiments about the covid-19 vaccination policy with concerns, among others, about vaccine safety issues and the potential for corruption from the policy. To increase public confidence in the covid-19 vaccination policy, the government must answer them. Concerns about vaccine safety can be overcome by giving an example of a high government official of being willing to be vaccinated beforehand and live broadcast as was done by the

elected president and vice president of the United States Joe Biden and Kamala Harris. Meanwhile, concerns over the potential for corruption can be resolved by taking more transparent, accountable and responsive actions in implementing the vaccination policy.

Keywords : Covid-19 Vaccination; Sentiment Analysis; Public Policy; Public Trust

Introduction

The COVID-19 pandemic that occurred around the world since 2020 has taught many countries in the world to take trial and error policies in dealing with this completely new situation. Many countries are considered to have failed in handling the COVID-19 pandemic, but there are also countries that are considered successful. One of them is South Korea, which is considered successful because it was able to reduce the number of infections without implementing a massive lockdown. One of the keys to South Korea's success in dealing with the COVID-19 pandemic is the openness of data and information that has been carried out since February 2020. It has been able to make the people of South Korea believe in the actions taken by the government in handling the pandemic. Through this, the people of South Korea voluntarily take preventive measures for the transmission of COVID-19 such as wearing masks, washing hands and self-quarantine, and even being actively involved in helping the government such as making applications to track COVID-19 patients (Moon, 2020).

Public trust in government policies is one of the successes in handling the COVID-19 pandemic. With public trust, the community will participate voluntarily in various policies taken by the government. The policy includes not only adherence to health protocols but also participation in vaccination programs. It is important because the success of vaccination program depends on the number of people being vaccinated. The more the people are vaccinated, the faster the herd immunity will be created. In August 2020, President Jokowi announced that the vaccination program would be implemented in January 2021. In December 2020, the Sinovac vaccine from China arrived at Indonesia. The availability of vaccines shows the government's readiness to run a vaccination program in early 2021. Various reactions from the community have emerged since then. Some people support and are willing to participate in vaccine clinical trials, but some others refuse vaccination for various reasons.

The reaction or public sentiment on vaccination policy, according to the writer, is very interesting to analyze. The community's positive sentiment indicates that there is support for the implementation of this program, while negative sentiment indicates the community's rejection. Public sentiment, either positive or negative, can be input in evaluating the vaccination program so that they can voluntarily participate. In this article, the writer will try to collect and analyze public sentiment on COVID-19 vaccination in Indonesia using data from social media, which is public comments on videos about Covid-19 vaccinations uploaded on YouTube. YouTube is the largest video platform today with 2 billion users and 1 billion hours of content watched every day

(YouTube, 2020). In Indonesia, the number of unique users reaches 93 million (over 18 years old) who watch YouTube videos every month (Tesalonica, 2020).

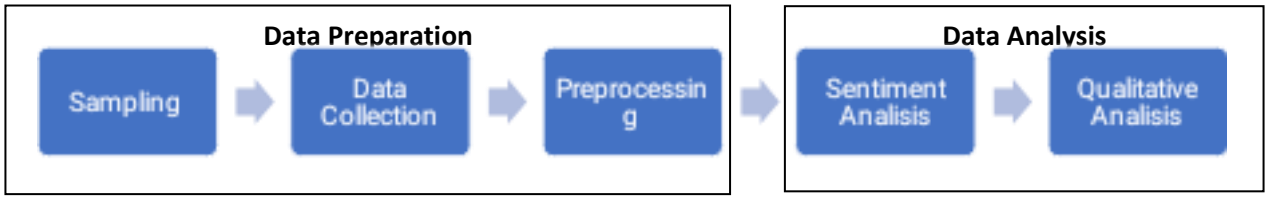
Sentiment analysis is a method of taking information from a text and dealing with the intended direction of the writer through the language used (Liu, 2015). Sentiment analysis will classify a text into negative or neutral sentiment. This analysis is usually used for a website-based application of products such as books, movies or restaurants (Razia Sulthana & Ramasamy, 2017). In recent years, sentiment analysis has also been used widely to analyze social media data to find out netizens' opinions on a topic, such as research on twitter user sentiment about climate change (Dahal et al., 2019) and research on YouTube user sentiment on school zoning policies (Dahal et al., 2019) Dahal et al., 2019) Anggraini & Tursina, 2019).

Since the COVID-19 pandemic, many studies have been carried out on social media data to determine the public's response to COVID-19. For example, researches on public responses on social media Twitter to covid-19 for the sake of infodemiology (Xue et al., 2019), on public responses to newspaper headlines about covid-19 to find out public sentiments and emotions (Aslam et al., 2020), and to seek insight from social media about COVID-19 toward mental health (Valdez et al., 2020). This research is expected to be able to describe how people's sentiments are towards the covid-19 vaccination program and what things are taken into consideration to participate or reject the program.

Method

This study used purposive-sampling to extract comments from selected YouTube videos considered to represent the problem of covid-19 vaccination. The comments having been extracted will be entered into the data preprocessing stage and then become a dataset for sentiment analysis to determine public sentiment towards the COVID-19 vaccination policy. Furthermore, a descriptive qualitative analysis would be carried out to find out what topics are taken into consideration by netizens in accepting or rejecting the covid-19 vaccination policy. The stages of the study can be seen in Figure 1.

Figure 1.
Research Stages



Source: Researchers

This study used purposive-sampling by taking sample comments from 4 YouTube videos related to the issue of covid-19 vaccination in Indonesia. These videos are considered to represent vaccination issues in cyberspace and are uploaded by independent news portals (not for or against

the government) so that the netizens' comments are expected to be more objective. The sample videos can be seen in Table 1. Preprocessing is a data cleaning stage consisting of lowercase, removing punctuation, stemming, and removing stop word (Ibnugraha et al., 2018); thus, it is ready for further analysis. The detailed steps are as follows:

a. Lowercase: converting capital letters to lowercase letters for easy identification and word grouping.
b. Removing Punctuation: removing punctuation marks such as question marks (?), exclamation points (!), periods and commas.
c. Stemming: changing a word into its root word. For example: *berjalan* into *jalan* and *menjalankan* into *jalan*.
d. Removing stop word: removing meaningless stop marks. For example: and, or, which, with.

Datasets that have been preprocessed would be manually labeled positive, negative and neutral according to the sentiment of each comment. Thereafter, word weighting was carried out using the TF-IDF (Term Frequency-Inverse Document Frequency) method by scoring the occurrence of words in the document to find out how important the word is (Syarifuddin, 2020). The labeled dataset would be divided into two parts: training data and testing data to be tested using the Naïve Bayes algorithm. This test was carried out to predict each label given so that the accuracy, precision and recall values would be obtained. Sentiment analysis testing was conducted using the RapidMiner application.

Datasets having been labeled and processed using the TF-IDF method would be grouped based on the same topic to analyze what topics are taken into consideration by netizens in accepting or rejecting the covid-19 vaccination policy. This qualitative analysis used 6 steps of thematic analysis from Braun & Clarke (2006): (1) recognizing keywords from the data, (2) creating initial codes, (3) searching for themes, (4) evaluating potential themes, (5) defining themes, and (6) reports.

Table 1.
YouTube Sample Video

No	Title	Video Description	Upload Date	Video Owner	Number of Comments	Link
1	*Jokowi: Vaksin Covid-19 Siap Januari 2021*	President Joko Widodo expressed his hope that the Covid-19 vaccine would be produced in January 2021.	11 August 2020	CNN Indonesia	1179	https://www.youtube.com/watch?v=PwwL10ZlE6g
2	*Ini Urutan Prioritas*	The Head of P2P of the Indonesian	19 October	Kompas TV	799	https://www.youtube.com/

	Penerima Vaksin Covid-19 di Indonesia	Ministry of Health, Achmad Yurianto, explained the order of recipients of the Covid-19 vaccine in Indonesia.	2020			watch?v=ENmvCS-0pdc
3	*Vaksin Covid-19 Tiba di Tanah Air*	President Jokowi confirmed that 1.2 million Covid-19 vaccines made by Sinovac had arrived at Indonesia.	6 December 2020	CNN Indonesia	1750	https://www.youtube.com/watch?v=JbxULCgqLPg
4	*Jokowi: Vaksin Covid-19 Gratis untuk Masyarakat!*	President Joko Widodo ensured that the COVID-19 vaccine would be free of charge for the public.	16 December 2020	Kompas TV	934	https://www.youtube.com/watch?v=63UjXrDjr-o

Results and Discussion

Data collection was carried out on 22-23 December 2020 using Facepager application. The comments taken were the top level ones, rather than replies to a comment. There are 2265 comments collected. The comment dataset was then preprocessed. The first step is lowercase, which is changing capital letters to lowercase letters as shown in table 2.

Table 2.
Lowercase Process

Sample Comment	***Lowercase***
URUTANNYA BAGUS NYA : 1. DPR 2.PEMERINTAH & KELUARGA 3.PETUGAS KESEHATAN JGN KLU UJI COBA RAKYAT DULU TP KLU UANG KRONI NYA DULU .	*urutannya bagus nya : 1. dpr 2.pemerintah & keluarga 3.petugas kesehatan jgn klu uji coba rakyat dulu tp klu uang kroni nya dulu.*

The next stage is the process of removing punctuation marks that can be seen in table 3.

Table 3.
Punctuation Removing Process

Sample Comment	***Remove Punctuation***

urutannya bagus nya : 1. dpr 2.pemerintah & keluarga 3.petugas kesehatan jgn klu uji coba rakyat dulu tp klu uang kroni nya dulu .	*urutannya bagus nya dpr pemerintah keluarga petugas kesehatan jgn klu uji coba rakyat dulu tp klu uang kroni nya dulu*

Then the stop word or conjunction is removed as can be seen in table 4.

Table 4.
Stop Word Removing Process

Sample Comment	***Remove Stop word***
urutannya bagus nya dpr pemerintah keluarga petugas kesehatan jgn klu uji coba rakyat dulu tp klu uang kroni nya dulu	*urutannya bagus dpr pemerintah keluarga petugas kesehatan uji coba rakyat uang kroni*

After removing the stop word, the word is changed to its root word as shown in table 5.

Table 5.
Stemming Process

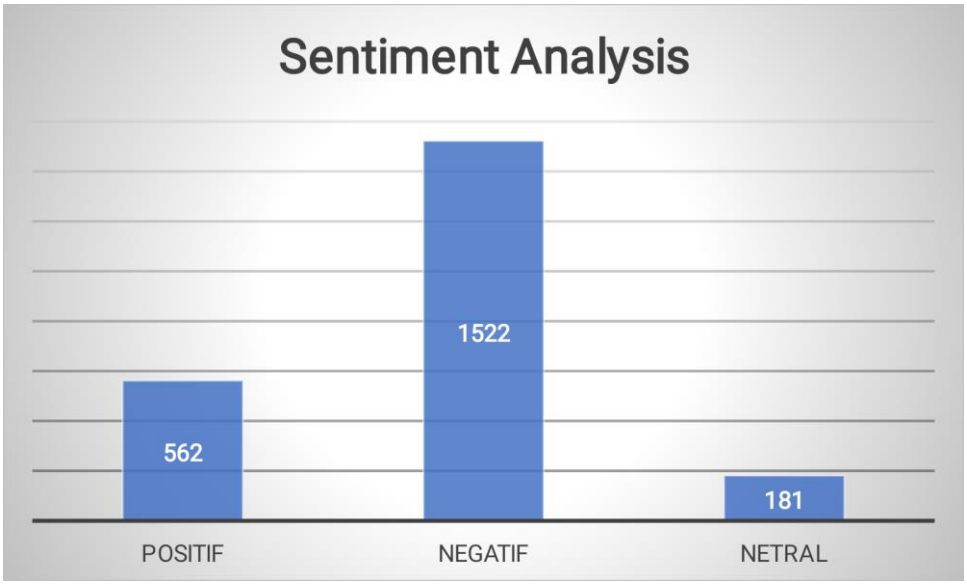
Sample Comment	***Stemming***
urutannya bagus dpr pemerintah keluarga petugas kesehatan uji coba rakyat uang kroni	*urutan bagus dpr pemerintah keluarga petugas kesehatan uji coba rakyat uang kroni*

The preprocessed dataset is then manually labeled for each comment: positive, negative or neutral. There are 562 comments with positive labels, 1522 with negative labels, and 181 with neutral labels as the result.

Table 6.
Examples of Labelled Comments

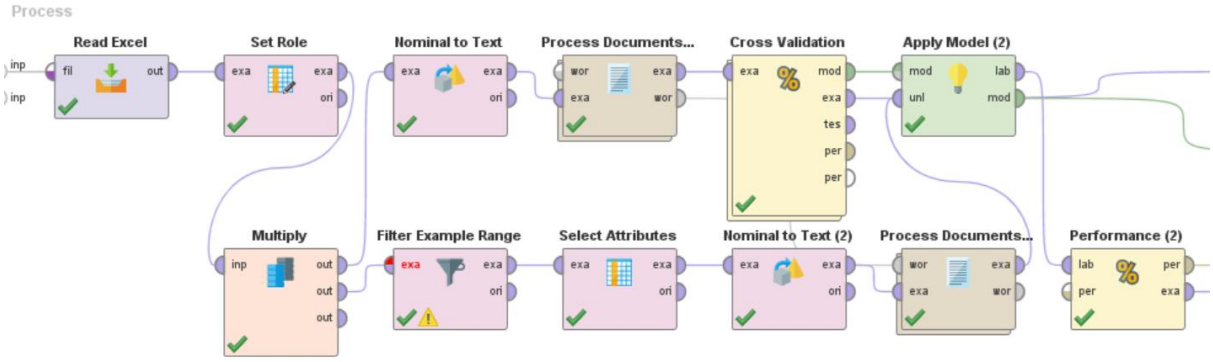
Comments	Label
alhamdulillah semoga vaksin usir virus covid 19 negara aman tentram amin	*positif*
tolak gratis waktu sd vaksin meriang seminggu bagaimana vaksin c19 ngeri	*negatif*
zona merah prioritas uji coba vaksin sinovac	*netral*

Figure 2.
Histogram of Sentiment Analysis Based on the Number of Comments



The labeled dataset will be tested using the Naïve Bayes algorithm. The dataset is divided into two parts: as training data and testing data, with a portion of 30% and 70% of the initial dataset. Then the dataset will be processed with the RapidMiner application as can be seen in Figure 3.

Figure 3.
The Testing Process on the RapidMiner Application



The test results provide a confidence value (positive, neutral, negative) for each label as shown in Figure 4 and an accuracy score of 91.32% is also obtained, as shown in Figure 5.

Figure 4.
Confidence Score Prediction Results for Each Label

Row No.	Polarity	prediction(P...	confidence(negatif)	confidence(positif)	confidence(netral)
1	negatif	negatif	1	0	0
2	negatif	negatif	1	0	0
3	positif	positif	0	1	0
4	netral	netral	0	0	1
5	netral	netral	0	0	1
6	netral	netral	0	0	1
7	negatif	negatif	1	0	0
8	positif	positif	0	1	0
9	positif	positif	0	1	0
10	netral	netral	0	0	1
11	positif	positif	0	1	0

Figure 5.
Naive Bayes Algorithm Accuracy Score

accuracy: 91.32%

	true negatif	true positif	true netral	class precision
pred. negatif	266	0	0	100.00%
pred. positif	10	253	0	96.20%
pred. netral	34	15	102	67.55%
class recall	85.81%	94.40%	100.00%	

The result of data processing using the TF-IDF method on the RapidMiner application shows the words and their number of occurrences in the dataset. These words are in the form of unigrams or single words, as well as bigrams or pairs of words. Bigrams are two words that appear in pairs regardless of grammar and semantics and sometimes cannot be directly understood (Xue et al., 2019). Unigrams that appear frequently include *"vaksin"*, *"semoga"*, *"covid"*, *"alhamdulillah"*, *"uji"*, *"tolak"*, while bigrams that appear include *"uji_coba"*, *"vaksin_covid"*, *"tolak_vaksin"*, *"alhamdulillah_semoga"*, *"semoga_vaksin"*. The top 50 unigrams and bigrams can be seen in table 7 and visualized in word cloud form in Figure 6.

Figure 6.
Wordcloud of Unigram/Bigram that Appears Frequently

Table 7.
Top 50 Unigrams/Bigrams

Words	Total Appearances	Words	Total Appearances
vaksin	941	*allah*	66
semoga	279	*pejabat*	66
covid	242	*vaksinnya*	65
rakyat	213	*vaksin_covid*	59
alhamdulillah	181	*takut*	57
indonesia	145	*mati*	56
coba	129	*bayar*	50
gratis	116	*menteri*	45
corona	105	*kena*	44
uji	104	*bisnis*	42
tolak	93	*tolak_vaksin*	42
virus	93	*beli*	40
negara	88	*bilang*	40
dpr	86	*aman*	39
china	80	*penyakit*	38

sehat	80	*alhamdulillah_semoga*	37
presiden	79	*kaya*	37
suntik	79	*koruptor*	37
uji_coba	76	*pandemi*	36
masyarakat	75	*dunia*	35
cepat	70	*normal*	35
korupsi	69	*sekolah*	34
pemerintah	69	*butuh*	33
amin	68	*januari*	33
duluan	68	*negeri*	33

Unigrams/bigrams and comments were analyzed to determine the classification of topics on the issue of vaccination. From the results obtained, there are 4 major groups of topics. They are: (1) support or optimism about vaccination policies, (2) distrust in the safety of the vaccines given, (3) fear of vaccination methods and (4) pessimism about vaccination policies because they are prone to corruption. The results of the analysis can be seen in table 8.

Table 8.
Results of Topic Analysis on Unigram/Bigram with Example Comments

Unigram/Bigram	Topics	Example Comments
semoga, alhamdulillah, alhamdulillah_semoga, amin, normal, sekolah, cepat, sehat	*Mendukung/optimis terhadap kebijakan vaksinasi*	*alhamdulillah bismillah semoga covid segera hilang amin*
uji, tolak, dpr, presiden, uji_coba, pemerintah, duluan, pejabat, mati, menteri, tolak_vaksin	*Tidak percaya terhadap keamanan vaksin*	*coba urutan pertama yang coba vaksin itu dpr menteri orang pemerintahan biar kalian lumpuh dulu kaya di luar negeri*
takut, suntik	*Takut terhadap metode vaksinasi yaitu berupa suntukan*	*gw tolak suntik takut jarum nya katanya pak rk gede ga bayang gw*
koruptor, korupsi, bisnis, kaya	*Tidak mendukung kebijakan vaksinasi karena rentan dikorupsi*	*ada aroma aroma korupsi gak tuh di pengadaan vaksin covid19 ini seperti di kementerian sebelah tu yg bantuan terdampak covid*

Closing

The Covid-19 vaccination policy is an important policy in tackling the pandemic because it is considered as the ultimate weapon to end this. However, from the analysis results of netizen comments on YouTube social media, it was found that the majority of netizens had negative sentiments towards the policy. The things being the concern of netizens are, among others, the issue of vaccination safety and the potential corruption.

The government must pay attention to the concerns of the public about the vaccination policy to improve the public's confidence in the policy and to enable the vaccination program to run smoothly. The concerns about vaccine safety issues can be overcome by setting the example of government officials willing to inject vaccines first as the president and vice president of the United States did. In addition, other sources of vaccine are needed besides Sinovac from China, because many netizens have negative sentiments against vaccines from China.

The government can answer public concerns about the potential corruption in the Covid-19 vaccination program by acting more transparently, accountable, and responsively, because these three things can prevent corruption and can affect public trust in the government.

References

Anggraini, N., & Tursina, M. J. (2019). Sentiment Analysis of School Zoning System on YouTube Social Media Using the K-Nearest Neighbor with Levenshtein Distance Algorithm. The 2019 7th International Conference on Cyber and IT Service Management, CITSM 2019.

Aslam, M.-F., Awan, T., Syed, J. H., Kashif, A., & Parveen, M. (2020). Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. Humanities and Social Sciences Communications, 7. https://doi.org/10.1057/s41599-020-0523-3

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3, 77–101. https://doi.org/10.1191/1478088706qp063oa

Dahal, B., Kumar, S. A. P., & Li, Z. (2019). Topic modeling and sentiment analysis of global climate change tweets. Social Network Analysis and Mining, 9(1), 24. https://doi.org/10.1007/s13278-019-0568-8

Ibnugraha, P. D., Nugroho, L. E., & Santosa, P. I. (2018). An approach for risk estimation in information security using text mining and jaccard method. Bulletin of Electrical Engineering and Informatics, 7(3), 393–399. https://doi.org/10.11591/eei.v7i3.847

Liu, B. (2015). Sentiment analysis: Mining opinions, sentiments, and emotions. Cambridge University Press.

Moon, M. J. (2020). Fighting COVID-19 with Agility, Transparency, and Participation: Wicked Policy Problems and New Governance Challenges. Public Administration Review, 80(4), 651–656. https://doi.org/https://doi.org/10.1111/puar.13214

Razia Sulthana, A., & Ramasamy, S. (2017). Context based classification of reviews using association rule mining, fuzzy logics and ontology. In Bulletin of Electrical Engineering

and Informatics (Vol. 6, Issue 3, pp. 250–255). Institute of Advanced Engineering and Science. https://doi.org/10.11591/eei.v6i3.682

Syarifuddinn, M. (2020). Analisis Sentimen Opini Publik Terhadap Efek Psbb Pada Twitter Dengan Algoritma Decision Tree,Knn, Dan Naïve Bayes. Inti Nusa Mandiri, 15(1). https://doi.org/10.33480/inti.v15i1.1433

Tesalonica. (2020). Jumlah pengguna unik YouTube di Indonesia capai 93 juta. Retrieved from Https://Www.Tek.Id/Tek/Jumlah-Pengguna-Unik-Youtube-Di-Indonesia-Capai-93-Juta-B1ZT79iPE.

Valdez, D., ten Thij, M., Bathina, K., Rutter, L. A., & Bollen, J. (2020). Social Media Insights Into US Mental Health During the COVID-19 Pandemic: Longitudinal Analysis of Twitter Data. J Med Internet Res, 22(12), e21418. https://doi.org/10.2196/21418

Xue, J., Chen, J., & Gelles, R. (2019). Using Data Mining Techniques to Examine Domestic Violence Topics on Twitter. Violence and Gender, 6(2), 105–114. https://doi.org/10.1089/vio.2017.0066

YouTube. (2020). YouTube dalam angka. Retrieved from Https://Www.Youtube.Com/Intl/Id/about/Press/.