Sentiment Analysis on the Indonesian Tweet of COVID-19 and COVID-19 Vaccination

Viskasari P. Kalanjati*[1], Nurina Hasanatuludhhiyah*[1], Annette d'Arqom[1], Ancah CN. Marchianti[2], Azlin Muhammad[3], Diana Purwitasari[4]

[1]Department of Anatomy, Histology and Pharmacology, Faculty of Medicine, Universitas Airlangga, Indonesia

[2]Department of Public Health, Faculty of Medicine, Universitas Airlangga, Indonesia

[3]Department of Parasitology and Entomology, Faculty of Medicine, National University of Malaysia

[4]Department of Electronic, FTEIC, Institute of Sepuluh Nopember, Indonesia

*Corresponding author: viskasari-p-k@fk.unair.ac.id

Professor in Department of Anatomy, Histology and Pharmacology, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia

Running title: Tweet analysis on COVID-19 and COVID-19 vaccination

Abstract

Objective: We analyze sentiments and opinions regarding Covid-19 and the Covid-19 vaccination on Indonesian-language Twitter. Apart from that, we also analyze fake news and facts as well as twitter engagement.

Materials and methods: We collected 3489367 data from March 2020 to August 2021. We analyzed fact and fake news using the string comparison method and used the difflib library to measure news similarity. We analyze engagement by averaging the engagement metrics of tweets, retweets, favorites, replies, and posts shared with sentiment and news opinion regarding Covid-19 and Covid-19 vaccinations.

Result: We found that public opinion on Twitter is consistent regarding Covid-19 that positive comments predominate. This is evidenced by the many tweets related to public enthusiasm regarding maintaining health, keeping distance, and handling to avoid the virus. Then, we found that positive and negative comments for public sentiment and opinion regarding the Covid-19 vaccination both dominate. Although people are enthusiastic about the vaccination program, there are also some people who are hesitant and even refuse to be vaccinated with certain types of vaccines. The refusal was due to various reasons such as side effects/KIPI or vaccine safety. Meanwhile, the results of twitter retweet engagement are more dominant, which proves that the public is enthusiastic in disseminating information related to Covid-19 and Covid-19 vaccination.

Conclusion: Overall understanding public sentiment and opinion can help related parties in handling the spread and prevention of the virus because each country handles it differently.

I. Introduction

In March 2020, the World Health Organization (WHO) declared COVID-19 a global pandemic. Due to the rapid spread, various preventive measures such as maintaining distance, wearing masks, and reducing mobility have been implemented. However, long-term pandemic prevention is needed, such as vaccination (Chou and Budenz, 2020). Various types of vaccines have been developed and approved in many countries to reduce the impact of the spread of the virus (Haque and Pant, 2020). However, multiple factors become obstacles in the implementation of vaccination, one of which is public doubts about the effectiveness and efficiency of vaccination programs to prevent COVID-19. Therefore, it is necessary to know public opinion about vaccination.

Various surveys have been conducted to find out public opinion about vaccination through social media communities, such as Twitter. Twitter has been used as a medium for discussion and argument on the spread of infectious diseases. In addition, Twitter data provides real-time public opinion on vaccinations compared to traditional surveys (Sinnenberg *et al.*, 2016). Various studies were conducted to identify on social media related to vaccination. (Chew and Eysenbach, 2010) identified rumors of a COVID-19 vaccine and conspiracy theories regarding vaccine effectiveness. (Du *et al.*, 2017) to know the public sentiment on social media towards vaccination. (Deden Ade Nurdeni, Indra Budi, 2021) monitors real-time public opinion regarding vaccine types. In addition, vaccination-related activities can affect the polarity of tweet counts (Tavoschi *et al.*, 2020).

Negative tweets are one factor in doubting vaccination (Featherstone *et al.*, 2020). Vaccine hesitancy is recognized as a global health threat (Lucía *et al.*, 2018). Various reasons for doubting vaccination arise, such as religious reasons, personal beliefs, and vaccine safety issues. Disinformation narratives spread on social media are sometimes hostile, causing anxiety, fear, and distrust of vaccination (Pulido *et al.*, 2020). Therefore, it is necessary to manage disinformation to increase vaccination acceptance in the community.

Security issues and government policies in various countries also affect vaccination activities (Ward *et al.*, 2020). Multiple studies have been conducted regarding public opinion regarding vaccination in multiple countries, as in the study (Subramanian and Kumar, 2021) that occurred in the US regarding the surge in COVID-19 cases due to low vaccination acceptance in the regions. Likewise, vaccination rejection caused a spike in COVID-19 instances (Velavan and Meyer, 2020) . Another study (Chou and Budenz, 2020)

researched Middle Income-Countries (MICs) with limited resources and capacity. Vaccination procurement is also influenced by politics, causing the vaccination process to be hampered.

Indonesia has the highest COVID-19 cases in Southeast Asia with high vaccination resistance (Chhetri *et al.*, 2020) . The comparison between the population and the distribution of vaccinations (Gao *et al.*, 2020) and public doubts are some of the obstacles to the vaccination program in Indonesia. The Ministry of Health of the Republic of Indonesia stated that a survey on vaccination by the COVID-19 Symptom Survey conducted by the University of Maryland Joint Survey Methodology Program in partnership with Facebook shows that several factors cause doubts about vaccination in Indonesia. Of the adults who are hesitant about vaccines in Indonesia, 49.2% are worried about side effects, and 34.9% want to wait and see the situation first as the main reason for doubting. Meanwhile, vaccine hesitancy in Indonesia varies the most between age groups from the main demographic groups. In particular, the youngest age group is the group most doubtful about vaccines, with the 18-24 year age group at 20.9% and the 25-34 year age group at 21.4% (Indonesia, 2021).

The presentation of the survey's result that has been carried out to the Indonesian people shows the high level of public doubt about the effectiveness and efficiency of the vaccination program. People's doubts arise due to misinformation because the narrative is sometimes harmful, causing anxiety, fear, and distrust, which leads to the rejection of the vaccination program. This study was conducted to find a public opinion about vaccination and Covid-19 on Indonesian-language Twitter text data. The research aims to analyze trending topics, find out tweets related to pro-vaccine and anti-vaccine, find out tweets related to vaccination categorized as fake news and facts, find out the public sentiment, and determine tweet exposure and engagement. The study results can be a way of making decisions by related parties to reduce disinformation which is a factor in people's doubts about the vaccination program.

II. Materials and Methods

As shown in Figure 1, several stages were carried out in this study to identify and analyze public opinion. The process starts from data preparation which consists of determining search terms for Covid-19 and Covid-19 vaccinations, then crawling Indonesian-language tweets based on the meta tagging language stored on Twitter and the search terms listed in table 1. The data first goes through the preprocessing stage. The data is cleaned of several elements such as emoticons, hashtags, non-alphanumeric characters, and URLs. The lower text is processed to convert all Twitter text to lowercase in the first step. The next step is to remove the punctuation in the Tweet text. The cleaned data that has gone through the preprocessing stage is then labeled manually. The data was labeled into two classifications, namely Covid-19 and Covid-19 vaccinations (type of vaccine, vaccine effectiveness, vaccine side effects). The labeling is then carried out

into several classes, namely fake and original tweets; Other courses are positive, neutral, and hesitate tweets. The analysis period begins January 1, 2020 - January 15, 2021, and January 16, 2021 - August 31, 2021. The total dataset obtained is 348,937 consisting of 24,579 data related to Covid-19 and 324,358 data related to Covid-19 vaccination.

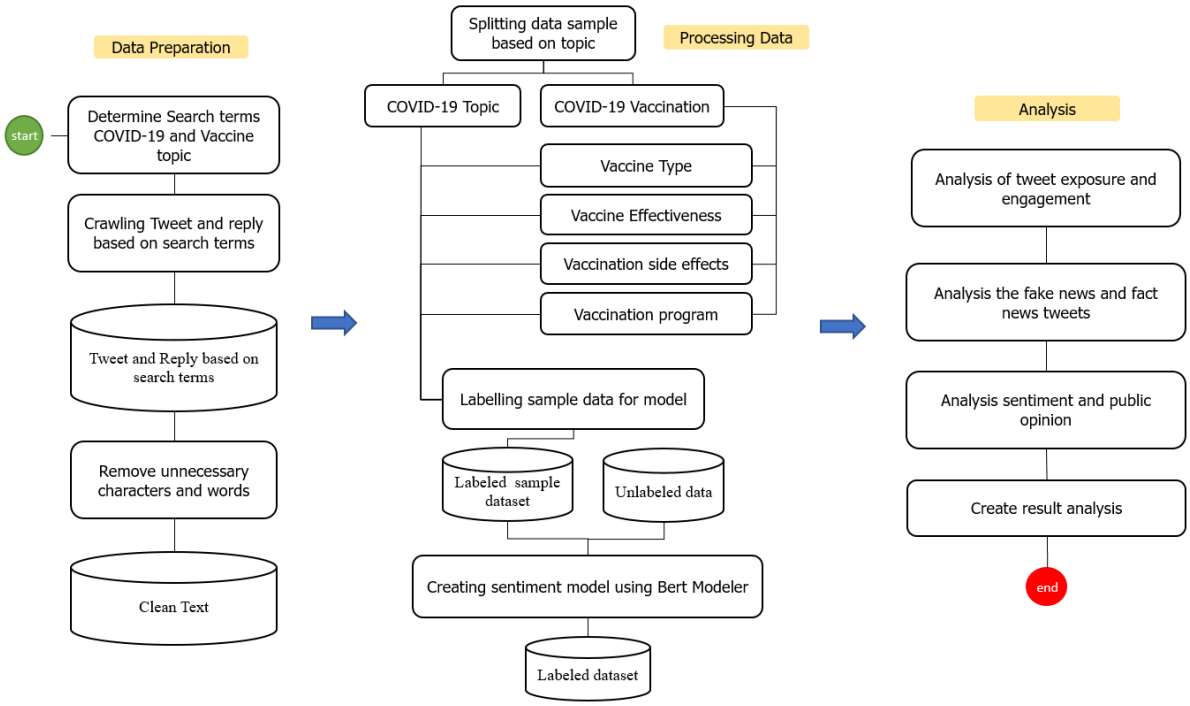

Figure 1. The process of dataset on the sentiment analysis on Indonesia Tweets during 1st of January, 2020 – 31st of August, 2021.

Table 1. Search terms COVID-19 and Vaccine COVID-19

Category	Search Terms
COVID-19	COVID-19, corona, corona virus, corona virus, SARS COV-2, COVID, masks, keep your distance, physical distancing, social distancing, washing hands, PPKM, PSBB, lock down, WFH, LFH, online learning, self-isolation, swab, PCR, 3M, 5M, 6M, tracing, comorbid COVID-19.
COVID-19 Vaccination	

Vaccine Type	COVID-19 vaccine, Sinovac, Chinese vaccine, Nusantara vaccine, red and white vaccine, biopharmaceutical vaccine, inactivated vaccine, mRNA vaccine, Zeneca Astra vaccine, Pfizer vaccine, Moderna vaccine.
Vaccine Effectiveness	Immunity, prevent COVID transmission, COVID positivity rate, positive COVID, herd immunity, death rate; delta variant, delta strain, variant of concern, prevent severe COVID, prevent ICU needs, prevent MRS needs.
Vaccine Side Effect	KIPI, post-immunization co-occurrence, paralysis, blood clots, blood clots, vaccine death, death after the vaccine, allergy, positive for COVID-19, drowsiness, hunger, sexual disorders, vaccine side effects, vaccine hazard, stroke, Guillain barre syndrome, pain, swelling, dizziness, headache, fever, muscle aches, vaccine chip.
Vaccination	National COVID vaccination, national COVID immunization, vaccination acceleration, mass vaccination, health worker vaccination, elderly vaccination, child and adolescent vaccination, third dose vaccination, booster, vaccination stage 1, vaccination stage 2, vaccination stage 3 vaccination, vaccination of BUMN, vaccination certificate, vaccination fake, self-vaccination, vaccination age 12-17, vaccination age 18-59, cooperation vaccination, paid vaccination, vaccination comorbid.

Search terms shown in Table 1 Post-Immunization Events (KIPI) is a form of the body's response to the injected vaccine. This KIPI is divided into two groups, namely mild and severe KIPI. After labeling, the next step is classification. However, before classifying according to a predetermined class, the topic of Covid-19 and vaccines was developed using the clustering method. The clustering method used is IndoBERT model by IndoNLU. After the clustering process is complete, the results of the data containing Covid-19 sentiments are continued with the data visualization design using python visualization. Detailed visualizations include pie charts, line charts, bar charts, tweet average engagement rate, and word cloud. Data visualization is used to determine public opinion. The test results will be analyzed in several stages, namely as follows:

1.1. Determination of the fake news and fact news tweets on COVID-19 and COVID-19 vaccination.

The first analysis determines fake news tweets and facts about Covid-19 and Covid-19 vaccinations. The string comparison method is used to assess fact or fake tweets. This method

compares tweets with a list of fake tweets in the list of fake tweets dictionary issued by TurnbackHoax.Id. However, this method has a weakness where if there is an element whose structure does not match the dictionary, it is considered a fact tweet, when in fact, this tweet contains aspects of fake news. The *difflib* library is used to get the similarity value between tweets. In determining tweets labeled fake news, it is necessary to have an additional parameter in the form of trace hold to tolerate the similarity of a tweet that can be considered a fake tweet and a tweet that is still considered a fact. The range of similarity values is determined based on a sample test by taking into account the data results with a range of 0 – 1, where the closer the value to 1 is, the more appropriate the word is in the list of incorrect tweets dictionary. After observing the tweet data, the trace hold value is determined to be 0.7, meaning that if there are tweets with a similarity value above 0.7, then the tweet is included in the category of fake tweets. From tweets that are indicated to be fake, we also try to search based on a dictionary list of tweets, including misleading content or fake content. The following are examples of tweets that are indicated as tweets with fake news:

Table 2. An example of a tweet that is indicated as a tweet with fake news

Username	Tweet	Content Classification
infocovid19_id	#HoaxBuster [SALAH] 21% Pasien Mengalami Efek Samping Setelah Memakai Vaksin Moderna Selengkapnya: https://t.co/Nh3DxlJIbr	fake news
cirtbuleleng	Dunia Setujui Vaksin Nusantara https://t.co/iWBlfFBTXM	fake news

1.2. Analysis of public opinion and sentiment on COVID-19 and COVID-19 vaccination before and after COVID-19 vaccination is started in Indonesia

The second analysis analyzes public opinion and sentiment on vaccination and Covid-19 before and after the Covid-19 vaccination began in Indonesia. Datasets from public and private accounts that have gone through the preprocessing stage are labeled based on positive, negative, and neutral criteria, as well as knowing the pattern of agreement on vaccination program data that have been marked as Pro Vaccine, Anti Vaccine, Doubtful, and Neutral which refers to the study (Yousefinaghani et al., 2021). Pro vaccines are categorized for tweets with a positive tendency towards

vaccination. Tweets show that the public can well accept the existence of vaccines and invitations to participate in vaccinations, even when giving an opinion about their condition after the vaccination process, commonly known as a follow-up event after immunization (AEFI). Anti-Vaccine is given in tweets that reject vaccination, accompanied by arguments against it. Doubt is given to tweets that tend to be confused about the purpose of vaccination or still doubt the effectiveness of certain vaccine brands, such as wanting only Pfizer vaccines and disparaging other brands of vaccines. Neutral is intended for tweets usually dominated by news accounts and only inform facts or narratives without expressing an opinion that says whether the statement is pro or anti-vaccine. A total of 3000 data were taken randomly from the dataset to be used as training data using the IndoBert model, with an accuracy rate of 75%. After the process is successfully carried out, we continue with data visualization and data analysis stages based on trends using python visualization.

1.3. Analysis of tweet exposure and engagement that discussed related to the COVID-19 and COVID-19 vaccination in Indonesia.

The final analysis analyzes tweet exposure and engagement that discusses Covid-19 and Covid-19 vaccinations in Indonesia. This section is used to measure the metadata that accompanies posts, including engagement measures such as the number of times each tweet was shared (retweeted), liked (favorite), people responded (replied), and retweeted with a comment (quote). Summary of engagement metrics calculated on average, including retweets, favorites, replies, and share of posts with positive or negative sentiment. The results of the engagement calculation continue with the stages of data visualization and data analysis based on trends using python visualization.

III. Result and Discussion

The rapid and complex spread of COVID-19 has made it known as a global pandemic, resulting in public unrest. The first COVID-19 case was announced to enter Indonesia on March 2, 2020 (Supriatna, 2020). There have been a lot of discussions about COVID-19, such as education and preventive measures by related parties and the public in all media, including Twitter. The graph shows that when it was announced to enter Indonesia at the beginning of the pandemic, the discussion about the pandemic was quite busy and remained consistent. However, there was an increase between May and July 2020. This was because at that time, Indonesia, which is the country with the largest Muslim population in the world, celebrated Ramadan and Eid al-Fitr during a pandemic which meant that the potential for the spread of COVID-19 massively increased. (Tfi, Hamblin and Rezaei, 2020). However, all health protocols are still being implemented, including the ban by the Indonesian government for all celebration activities during

Ramadan until Eid al-Fitr by taking social distancing (LSSD) measures. Even though national homecoming has been banned, local homecoming is still allowed. (Veruswati *et al.*, 2020). This has led to increased COVID-19 cases, followed by the discussion also growing in all media, including Twitter.

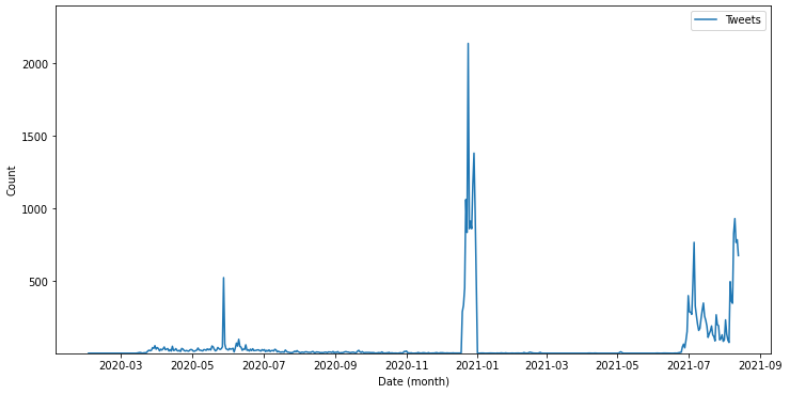

Figure 2. Tweets distribution on COVID-19 topic

The discussion of COVID-19 increased again at the end of 2020, followed by the Christmas and New Year holidays. The graph shows a significant increase in the debate of COVID-19 occurring. The analysis results show that the community carries out discussions about high mobility, and invitations and appeals to continue applying health protocols are quite a lot. Such as preventive measures to avoid the spread of the COVID-19 virus, namely staying at home, avoiding direct contact with other people, avoiding non-essential travel, social distancing, frequent hand washing, and so on (Tfi, Hamblin and Rezaei, 2020). Based on the data that has been collected by activity, the conversation about COVID-19 has started to decrease (not 0) until it rose again, recorded in the period after 2021. This coincided with the increase in attacks of the second wave of COVID-19 cases in July, where more and more people were exposed and full hospital bed occupancy.

Figure 3. Wordcloud COVID-19 *topic*

The Indonesian government has finally implemented the Emergency Public Activity Restrictions (PPKM) policy, attracting various comments from Twitter netizens. It has become a topic of discussion until the end of the data frame. As is known, the government implemented the PPKM program to reduce the spread of COVID-19. In addition, the direction of the Minister of Home Affairs regulates various things, including strengthening the implementation of testing, tracing, and treatment (3T) in each region. (Miharja *et al.*, 2021). The enactment of the policy brings positive and negative impacts, followed by the pros and cons of the surrounding community. The positive effect has been on reducing the rate of population mobility, reducing the Bed Occupancy Rate (BOR), and reducing red zones in several regions in Indonesia (Saraswati, Muzdalifah and Herawati, 2021). Meanwhile, the negative impact that the community has is that the implementation of the policy has made some people protest because they feel that the policy has an unfavorable effect on their economy. According to the analysis, discussions about PPKM are often a topic among the community. From the community activity sector that enforces mobility regulations related to PPKM, several sectors are of public concern, namely closing the purchasing center at 20.00 WIB / WITA / WIT and making the visitor capacity a maximum of 50% (Saraswati, Muzdalifah and Herawati, 2021). This causes a decrease in the income earned by some formal workers.

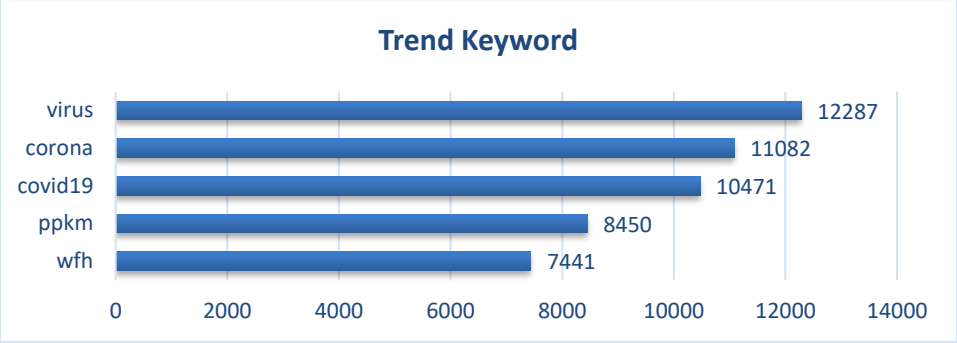

Figure 4. Trends in Covid-19 Conversation Attention before and after January 16, 2021

The word count graph shows many Twitter users discussed the COVID-19 virus along with tweets containing education to prevent exposure to the virus. During that period, the topic of discussion was still about education to suppress the spread of COVID-19. Preventive actions still being promoted, such as staying at home, avoiding direct contact with other people, avoiding non-essential travel, social distancing, frequent hand washing, and so on (Tfi, Hamblin and Rezaei, 2020) remain hot topics among the public. Meanwhile, the word count shows a hot topic among the people after implementing PPKM. After the government established and implemented the PPKM policy, Twitter users discussed correlated issues such as working from home. PPKM is intended as a form of response to the increase in COVID-19 cases, so the

problem of the virus is still quite busy being discussed on Twitter. Implementing PPKM levels 1-4 by the government, which has brought pros and cons to the community, has also become a topic of discussion. Even so, this policy is considered effective in suppressing the surge in the increase in COVID-19 cases (Saraswati, Muzdalifah and Herawati, 2021).

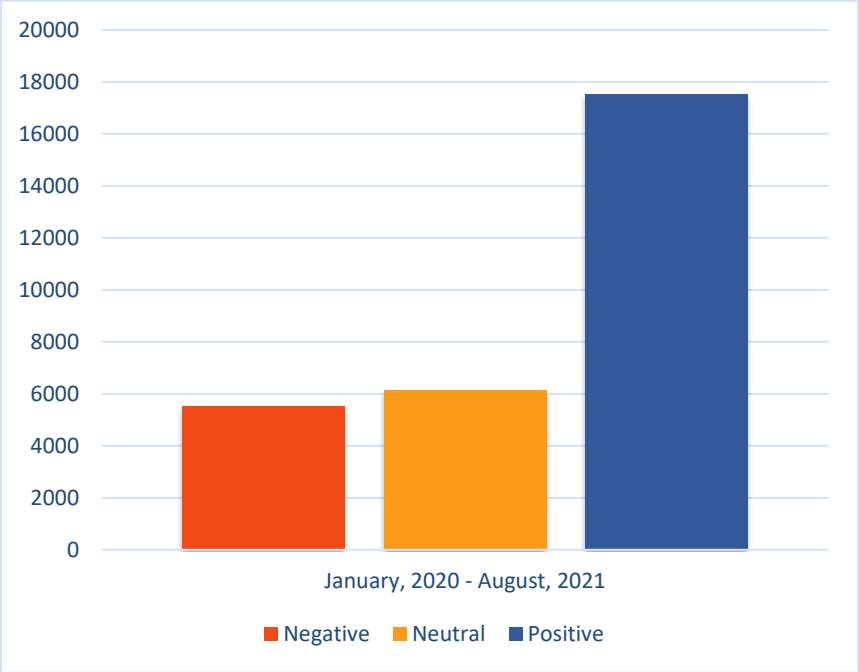

Figure 5. Public Sentiment about Covid Topic

The figure's visualization results explain where tweets with positive labels dominated the entire label. The sentiment referred to in the discussion includes users agreeing and understanding the conditions for COVID-19. The debate regarding the support for preventive actions by the government is also a topic that is still hotly discussed. In addition, Twitter users continue to carry out their activities as usual and provide education on how to prevent exposure to the virus, increasing positive sentiment about the covid issue. In the second period, the trend of positive sentiment led to discussions around the expressions of Twitter users to express their response to the increasing number of COVID-19 cases in Indonesia, accompanied by campaigns from all parties to carry out vaccinations aggressively. However, there was an increase in negative sentiment in the second period compared to the first period. This is because this period is when the government begins to issue PPKM policies that reap the pros and cons of the community. PPKM, which has a level of 1-4, is considered to harm the community's economy because of the limited activities of the community at work. This is because the rules made by the government are closing the purchasing center at 20.00 WIB / WITA / WIT and making the visitor capacity a maximum of 50% (Miharja *et al.*, 2021). Thus,

many parties have complained about the condition of the policy. If you visualize the sentiment distribution every month, it will be presented as shown in the following image.

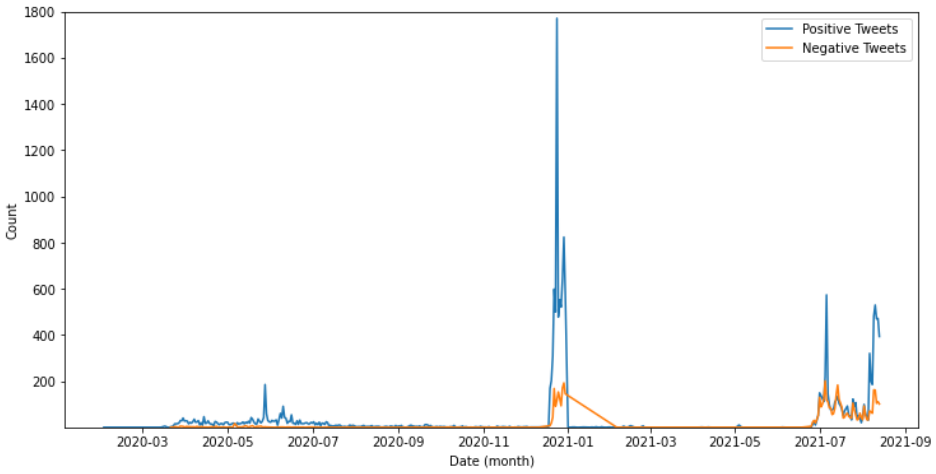

Figure 6. Covid Sentiment Distribution

Engangement COVID-19

Figure 7. Engagement COVID-19

In general, retweets dominate engagement, followed by replies. This supports evidence that the public is more focused on sharing information such as the rise or fall of cases as well as information on education and preventive measures implemented by the government to suppress the spread of COVID-19 (Hasanah, Suciati and Purwitasari, 2021). So, retweets dominate more than replies and likes. Meanwhile, replies are

in second place after retweets. Most people will ask questions and exchange information with other users through the reply column on every post related to COVID-19 to get more information.

Topic of Vaccine

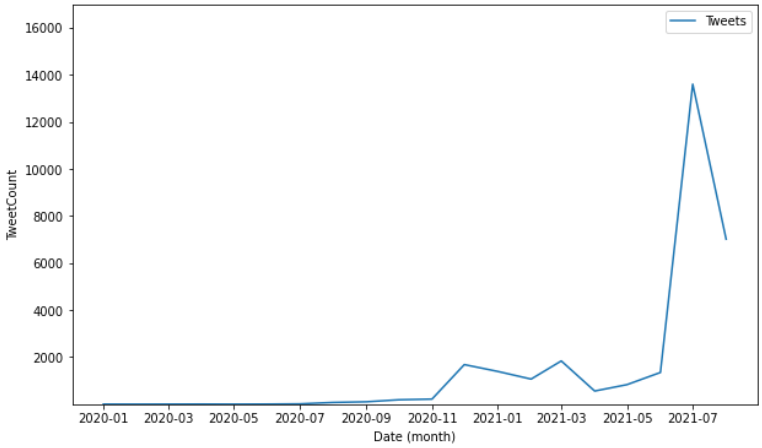

Figure 8 Tweet distribution about vaccine topic

Table 3. Vaccine topic tweets during 1st of January, 2020 – 31st of August, 2021.

Month	Count (Tweets)	
	2020	2021
January	1	1392
February	0	1066
March	2	1837
April	4	558
May	1	830
June	4	1347
July	13	13599
Agustus	73	7015
September	101	N/A
October	186	N/A
November	216	N/A
December	1682	N/A

In Figure 8 the discussion regarding vaccination began to increase in late 2020 and continued to increase until it reached its peak in July 2021. Vaccine products are starting to show hope in helping the community fight the coronavirus. In January 2021, Indonesia carried out a vaccination program so that various public opinions have emerged. The average number of tweets about vaccines based on the search keywords list starts from January until August 2020, with twelve tweets per month. The number is normal

because, during this period, the COVID-19 vaccine is still in the development stage by scientific experts. So that various community views respond to this vaccine development, such as the hope that a vaccine will be found soon to the community's negative response as if they think that making a covid vaccine is too long. Starting in September, we saw an increase in vaccine-related tweets because many prospective vaccine producers have shown good progress in developing their vaccines. As seen in Table 3, the number of tweets began to be discussed at the end of 2020 because the Sinovac vaccine was first brought to Indonesia on December 6, 2021. The discussion peaked in August when Indonesia was hit by a second wave of increased COVID-19 cases in all regions. The trend of vaccine tweets will continue throughout 2021 because the Indonesian government also prioritizes the community, especially the elderly, to vaccinate immediately. Then the vaccination program from the government was also getting more intense, so the discussion on vaccination grew. It is known, since the first case of COVID-19 was announced to infect Indonesia, discussions regarding preventive measures to prevent the spread of COVID-19 have become the most frequently discussed topics, one of which is about vaccines.To date, several clinical trials have been launched to test the effect of various vaccines against SARS-CoV-2 (Tfi, Hamblin and Rezaei, 2020). Like the vaccines that have been launched in Indonesia, namely Sinovac, AstraZeneca, Pfizer, and Moderna, which received various responses and comments from the public through Twitter social media, discussions about vaccines have reached more than 10,000 discussions and reached a top peak around October. At that time was the announcement of the efficacy of vaccines from various vaccine manufacturers.

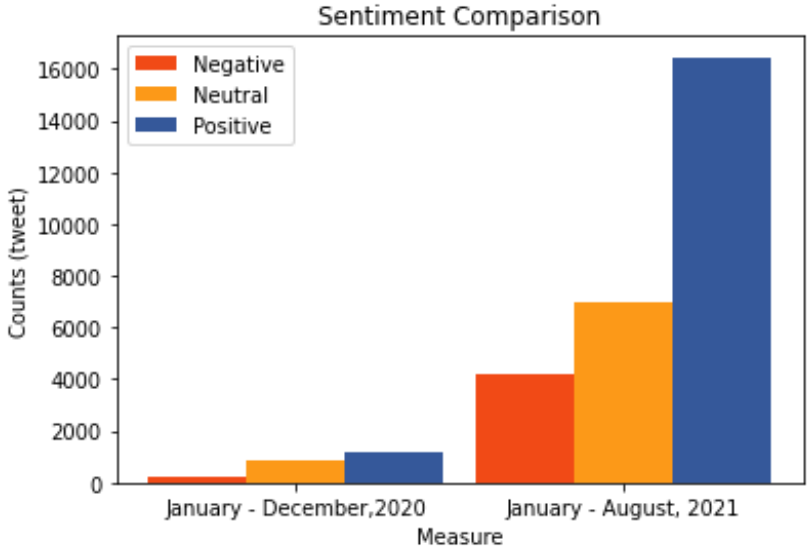

Figure 9. Vaccine Sentiment Analysis Comparison between before and after January 16, 2021

Sentiment analysis was conducted to find out the positive and negative sentiments of the public towards the vaccination program in the period before January 16, 2021, shows that positive tweet sentiment dominates

all existing sentiments. Analysis of the dataset indicates that generally, tweets come from news accounts where tweets originating from these accounts are classified as neutral tweets. While the period after January 16, 2021, shows that although the number of tweets is more dominant than in the first period, it shows that public opinion has a positive sentiment tendency. This is because the public is aware of the importance of using vaccines to help reduce the spread of the COVID-19 virus by participating in the vaccination program. However, negative sentiment is also quite dominating because people expect only certain types of vaccines or are still doubtful and do not believe in vaccination programs in combating the pandemic.

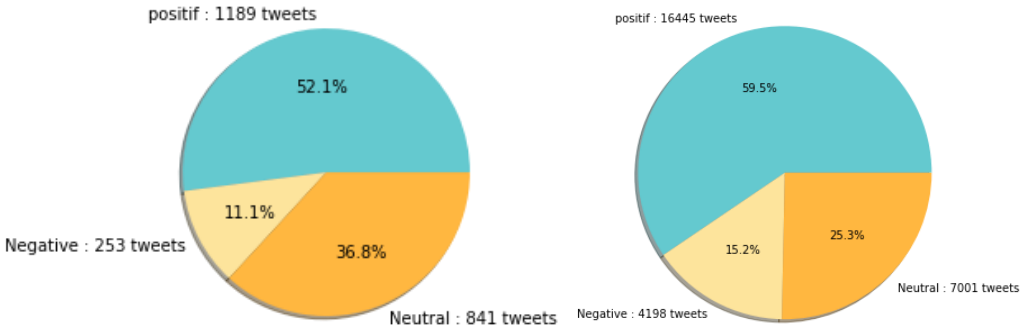

Figure 10 From the left sentiment of vaccine topic in 2020 and 2021

Table 4. Vaccine topic tweet analysis sentiment

Sentiment	2020			2021		
	Positive	Negative	Neutral	Positive	Negative	Neutral
January	0	0	1	822	121	449
February	0	0	0	600	59	407
March	2	0	0	1076	45	716
April	2	2	0	288	34	236
May	0	1	0	467	81	282
June	2	0	2	785	140	422
July	9	1	3	7617	3041	2941
August	21	5	47	4790	677	1548
September	34	5	62	N/A	N/A	N/A
October	104	17	65	N/A	N/A	N/A
November	144	15	57	N/A	N/A	N/A
December	871	207	604	N/A	N/A	N/A
TOTAL	1189	253	841	16445	4198	7001

The following table is in 2021, showing that positive Tweet sentiment tends to dominate across all timestamps. This proves that the public accepted it well during this period by providing positive feedback on all news of COVID-19, from the rise and fall of cases to information about several government policies

issued to prevent the explosion of COVID-19 cases in Indonesia. The previously mentioned preventive measures (Tfi, Hamblin and Rezaei, 2020) also received a positive response from the general public. This shows that the community supports the government's efforts in dealing with COVID-19 cases in Indonesia.

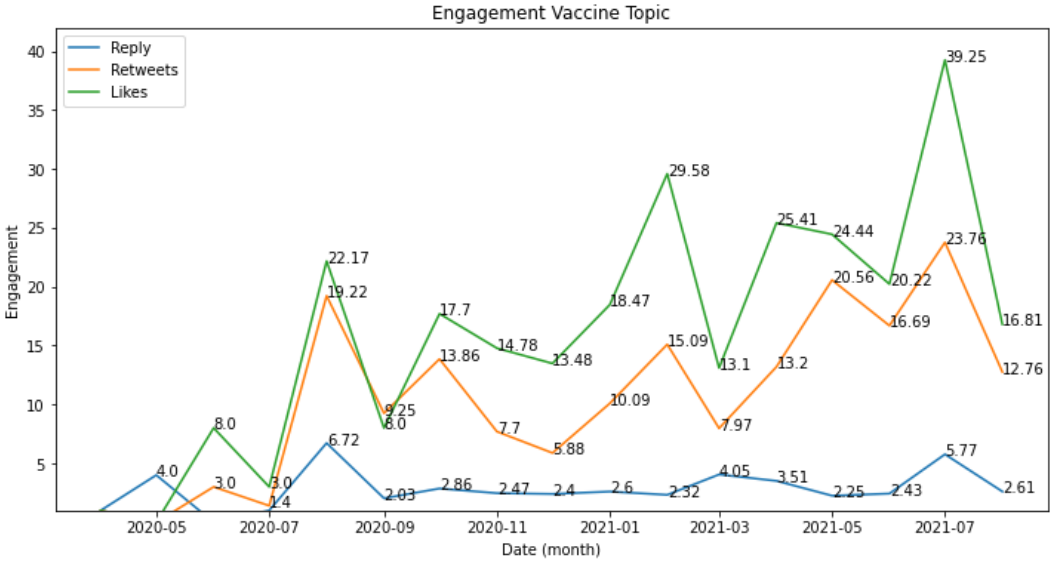

Figure 11 Engagement ratio vaccine topic

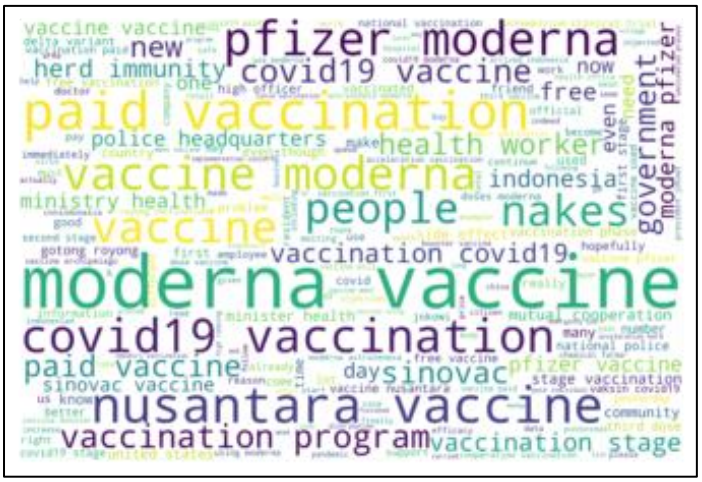

Figure 12 Wordcloud of vaccine topic

Table 5. Sentiment Based On Vaccine Brand

Time		Sentiment Based On Vaccine Brand				
		AstraZeneca	Sinovac	Moderna	Pfizer	Nusantara
2020	Positive	42	132	438	267	3
	Neutral	46	92	322	177	2
	Negative	9	27	100	62	1
2021	Positive	798	1697	7457	2344	2220
	Neutral	539	751	3348	1031	571
	Negative	185	420	1576	707	806

The table above shows the sentiments of Twitter residents regarding the type of vaccine launched to suppress the spread of COVID-19. There are vaccine brands that have been found from various countries in the world through research procedures that have been carried out (Tfi, Hamblin and Rezaei, 2020). Five types of vaccines were analyzed, namely AstraZeneca, Moderna, Pfizer, Sinopharm and Sinovac (Nasir *et al.*, 2021)(Bralianti and Akbar, 2021). In Indonesia, several brands that have gone through safe and halal tests that adapt to the conditions of Indonesia as a country with the largest Muslim population in the world are exciting topics to be discussed by the public. For example, the most popular brand in Indonesia is Sinovac. This is because Sinovac is the fastest vaccine brand to enter Indonesia. It can be seen from the data that has been successfully presented that Sinovac continues to dominate, especially at the end of 2020 and early 2021, when the Sinovac vaccine has entered Indonesia. Moderna vaccine also experienced an increase in the number of tweets because, during this period, many of these vaccines were spread in Indonesia. Moderna vaccine has the most different side effects from other vaccines, so conversation is increasing public. In addition, the AstraZeneca vaccine carries more significant side effects when compared to the Sinovac vaccine.

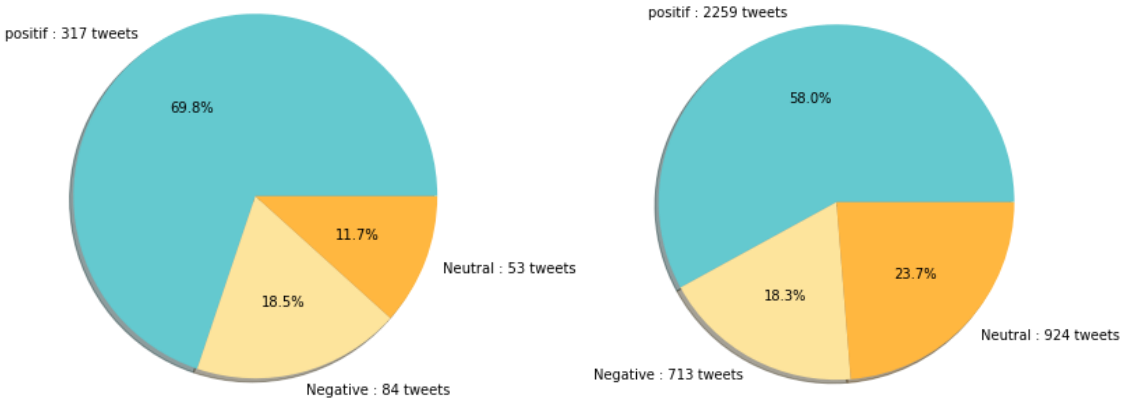

Figure 13 From the left, sentiment side effect vaccine topic in 2020 and 2021

The following chart, both in 2020 and 2021, positive sentiment tends to dominate above 50%. However, negative sentiment tends to be higher in 2020 when compared to neutral sentiment. Negative sentiment can happen because, in 2020, the existing vaccine research is still in the development stage by scientists. While in 2021, various vaccine products and study results and side effects are written and have gone through various stages of trials submitted by vaccine manufacturers have begun to be released. Due to the increasing clarity of reports regarding the impact and how the vaccine can speed up stimulating the body's immunity to the virus, the public has more confidence in the role of vaccines in accelerating recovery from the pandemic.

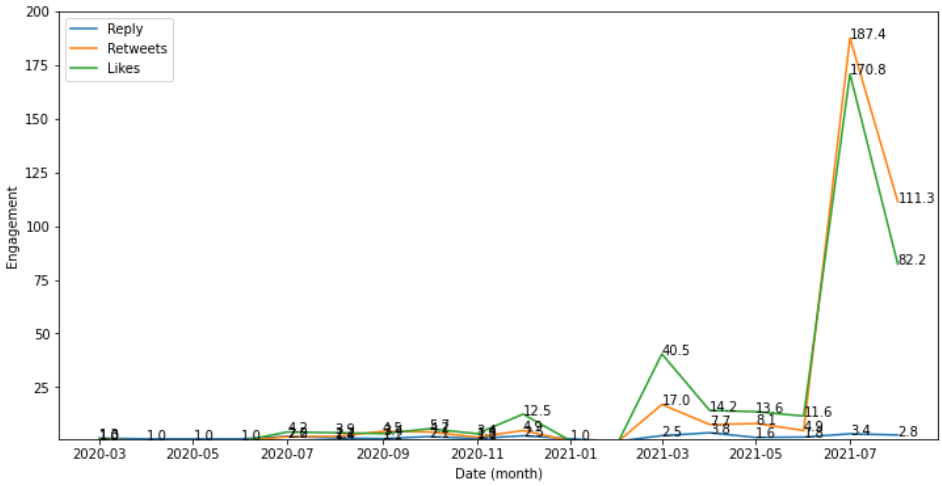

Figure 14 Engagement ratio of side effect vaccine topic

Meanwhile, figure 13 shows the average Twitter sentiment engagement per month, which shows that all the parameters as likes and retweets have more engagement from the distribution of existing data. Where in July 2021, the average retweets of existing tweets reached 187.41 retweets per tweet. An increasing tweet in July 2021 happened because two massive events took place this month. The first was that the vaccination program for the public had been carried out, and the government was aggressively carrying out mass vaccination activities. In contrast, another incident of Covid cases in Indonesia has experienced a reasonably extreme increase since June. The engagement ratio of the side-effect sub-topics that have been previously defined by ignoring tweets that do not have any engagement.

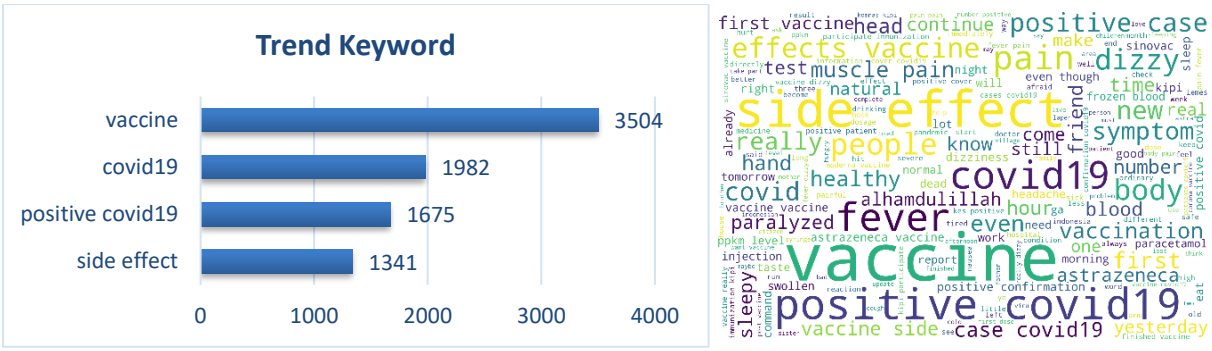

Figure 15. Wordcloud of vaccine side effects subtopics

The graphs and word cloud show that of the vaccine side effects subtopics recorded, the word most frequently discussed in tweets was vaccine topic. These words show results relevant to the keywords of this subtopic that are positive for COVID-19. Information about the rise and fall of COVID-19 cases and education to suppress the spread of COVID-19 are still words that often appear in discussions of COVID-19 by the public via Twitter.

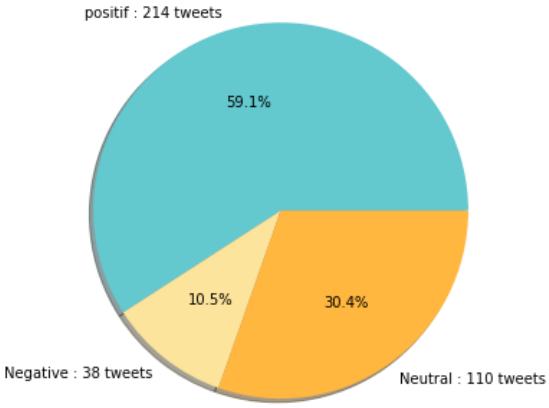

Figure 16. Sentiment of the effectiveness of the vaccination program

The following pie chart shows Twitter social media activities by the community regarding the effectiveness of the vaccination program. Positive sentiment dominates as much as 59.1% of the total sentiment, and only 38 tweets or about 10.5% of tweets are labeled as negative responses from the general public. Vaccination programs ranging from vaccine production to vaccine distribution are exciting topics discussed and want to be known more by the public. Phase 2/3 clinical trial results of nine vaccine candidates and implications for vaccine use and distribution in Indonesia (Ophinni *et al.*, 2020). The COVID-19 vaccine situation is moving very fast with new developments in a matter of days.

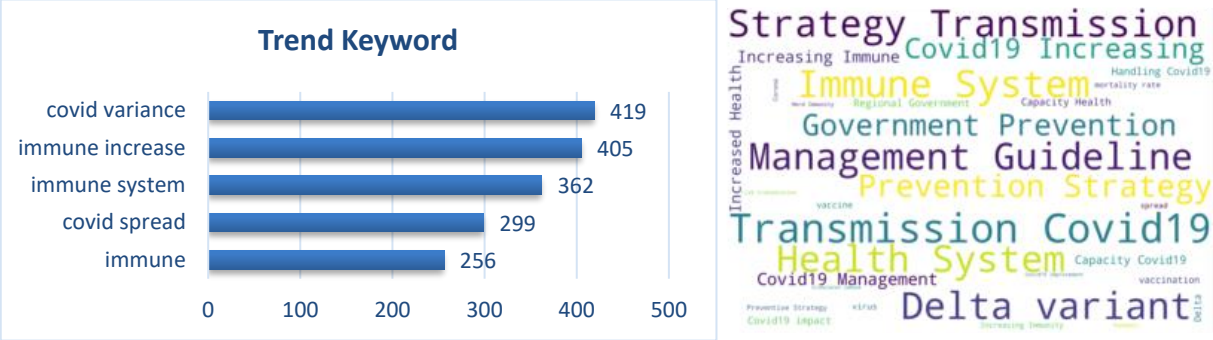

Figure 17. Wordcloud of the vaccine effectiveness subtopic based on keywords

The following graph shows that COVDI-19 still ranks first from the most discussed in the vaccine effectiveness subtopic based on keywords collected in a dataset with 900 tweets related to COVID-19. An increase follows this in immunity and the immune system, which is the goal of the vaccination program. Handling the spread of COVID-19, strategies to control the spread of the virus, the immune system, to the delta variant, which is the latest COVID-19 variant, are still topics that are often discussed in the case of vaccines. The effectiveness of vaccination is the topic of conversation that is most often addressed by the public following some information regarding public doubts about the efficacy of vaccination.

The type of Vaccine

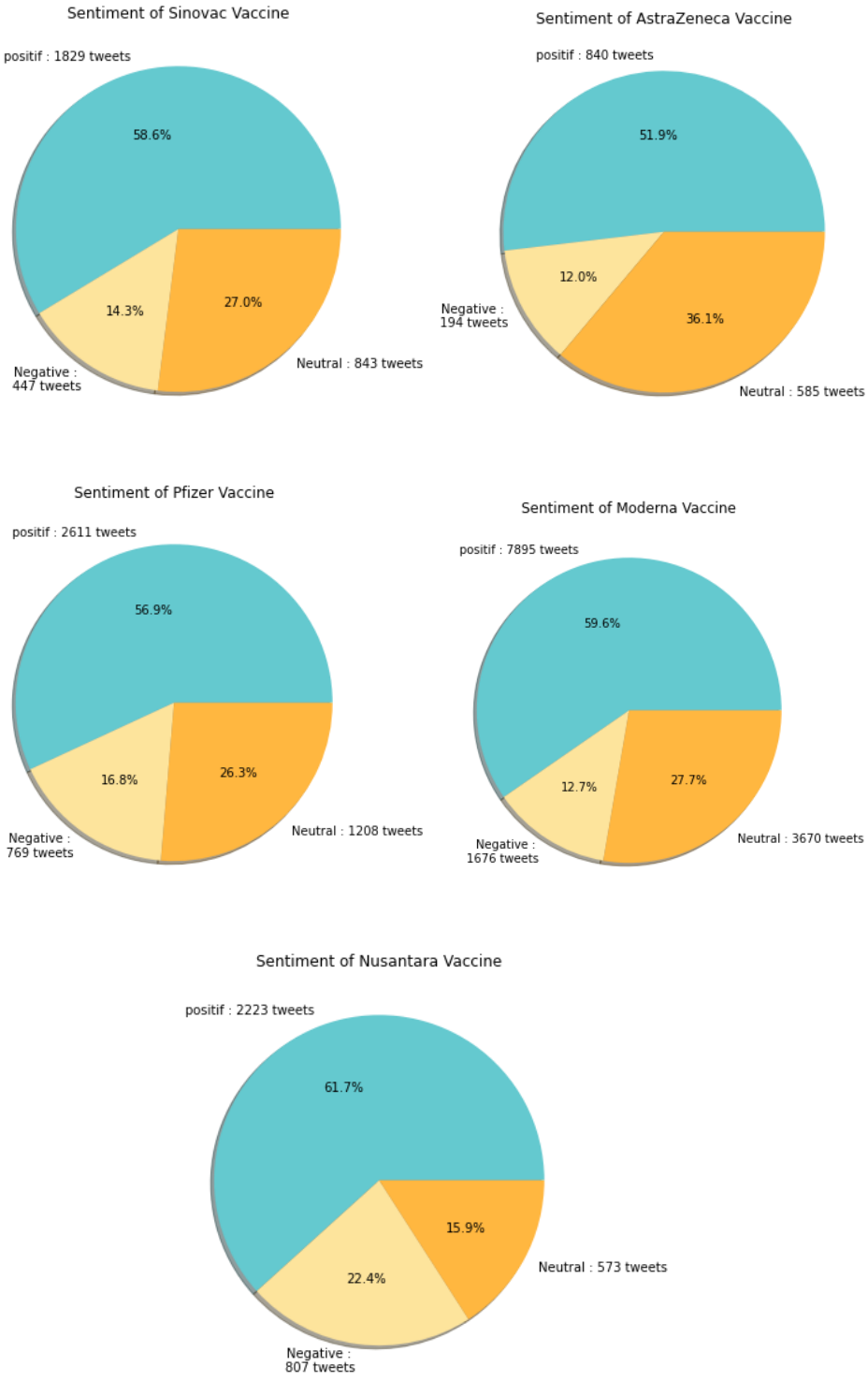

Figure 18 Sentiment Regards on Vaccine Brand

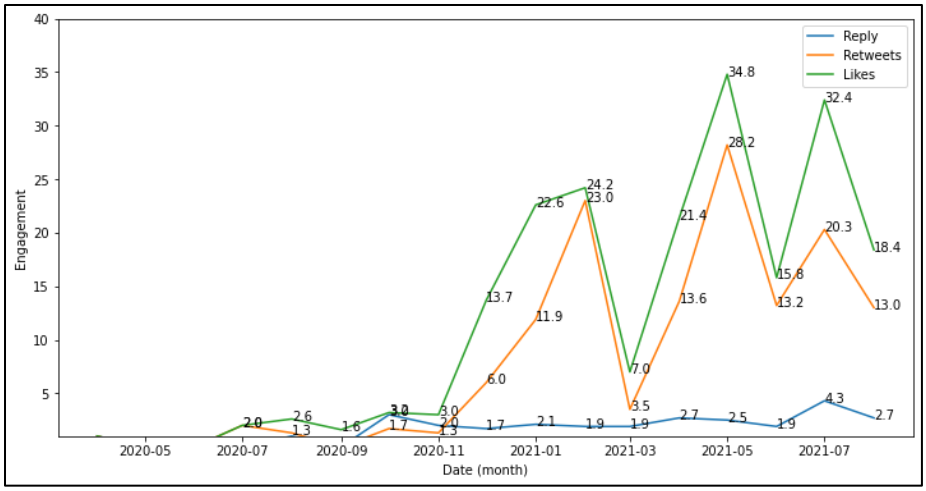

Figure 18. The type of Vaccine

In figure 17, positive sentiment dominates in all vaccine brands and is followed by neutral sentiment. Compared to the number of tweets discussing the COVID-19 vaccine from various manufacturers, there will be much discussion in 2021. The Moderna vaccine with a positive sentiment ratio of 59.6% is the vaccine that has been discussed the most. This is also influenced by the side effects that are so large compared to other vaccines but have a high level of effectiveness so that this vaccine is highly prioritized for health workers. Based on the figure 18, likes to tend to dominate among all. Peaked in May 2021, where the average likes were 34 likes of the collected tweets. This is because the public wants to share information about the types of vaccines disseminated to the public to overcome the spread of COVID-19. The many types of vaccines circulating make the people continue to seek information to study the functions and side effects of the vaccines received. Retweets tend to dominate, followed by likes, and the last is comments.

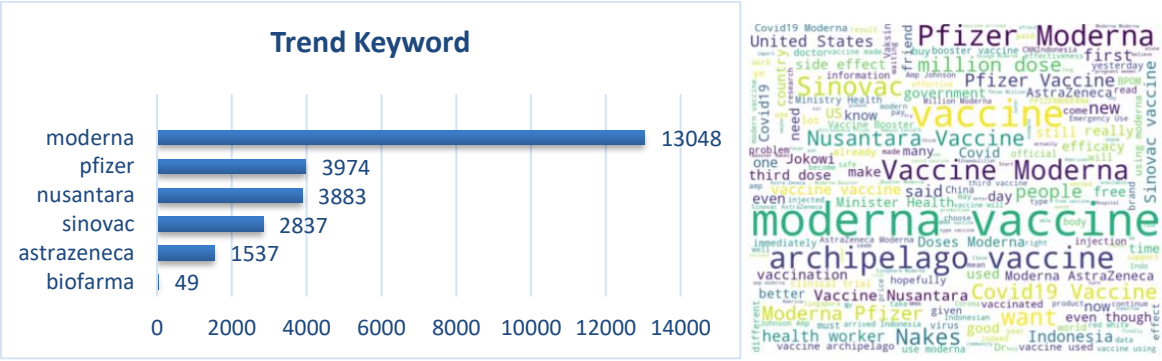

Figure 19. Wordcloud of the type of vaccine

Moderna is more often discussed in the subtopic of vaccine types, slightly different from the explanation at the beginning because Moderna here is limited to one sub-topic. At the same time, the

previous discussion has a broader scope, so that Sinovac is superior. It is known, Moderna is the latest type of vaccine that is disseminated to the public. Moderna is claimed to be a vaccine with superior effectiveness than the previous vaccines, namely Sinovac and AstraZeneca, first given to the general public (Ophinni *et al.*, 2020).

Vaccine program

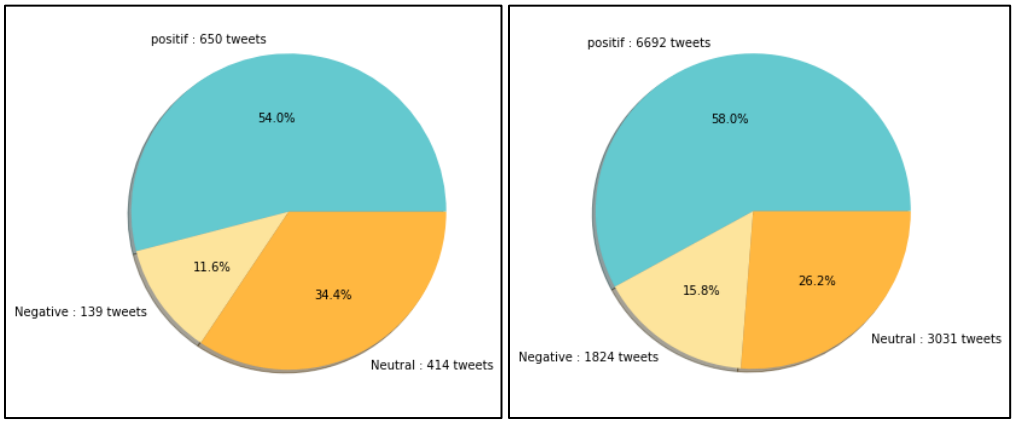

Figure 19 From the left sentiment of vaccine program in 2020 and 2021

In figure 20, both in 2020 and 2021 both have the same sentiment tendency, which is dominated by positive sentiment, although they are similar. However, both have differences in the quantity of discussion because, in 2020, there is no COVID-19 vaccination program that can be done. Public opinion began to increase in 2021 in response to various policies implemented by the government, for example, the public's pros and cons on administering vaccines that were subject to fees.

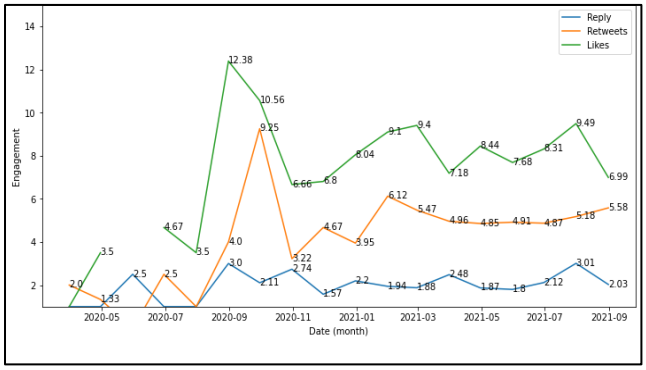

Figure 21. Sentiment and engagement ratio of the vaccination program topic

Meanwhile, in the discussion of the vaccination sub-topic, the average engagement shows that likes to dominate the timestamp, which was captured with a peak of 32 likes for one tweet. Meanwhile, in the second and third positions, namely retweet and reply. This shows that public interest in information about

vaccination is relatively high. This is evidenced by every tweet about vaccination getting a lot of likes. The public tends to like every tweet related to information about COVID-19. This is because they want to get and share information about COVID-19.

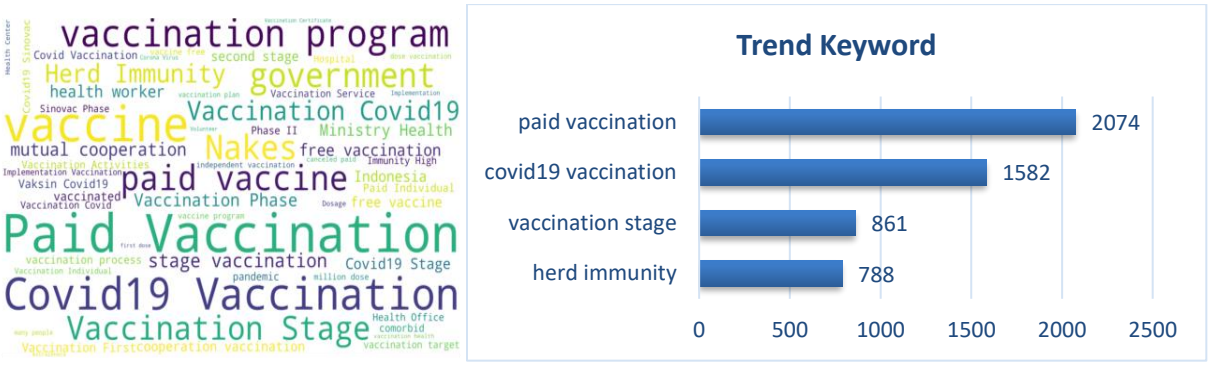

Figure 22. Wordcloud of the vaccination sub-topic

Regarding the sub-topic of vaccination, the most frequently discussed words are vaccination and vaccine. This has arisen from various issues that have emerged since the vaccination program. Some information states that community vaccinations are subject to different costs, differing from type to type. This makes the public's attention to all information about vaccination relatively high. The people also tend to express their opinion about the news with negative sentiments because they disagree with the issue of paid vaccination during a press conference for the government of the Republic of Indonesia.

Fake News Detection

Determination of true and false tweets uses the help of a dictionary that contains incorrect data on a news report in the community. Based on the results of observations of all tweets used in this study, it was found that 15 tweets were indicated as fake tweets with the classification of listed content as misleading content. The indicated content is about the world's approval of the indonesian vaccine, which in fact has not been tested until the last clinical trial. There is also content that contains relationships between several vaccine manufacturers affiliated with some companies. Another topic is that up to 21% of vaccine participants experience side effects after receiving the moderna vaccine. And the last one is the issue of vaccine manufacturing companies developing covid vaccines even before the covid-19 pandemic.

Tweets classified as fake tweets are dominated by the purpose of redistributing hoax information to the public which has been validated by Indonesia's Minister of Communication and Information. It is intended that the public better understand the facts regarding this fake news, as shown in the table. So, from the entire dataset, it can be concluded that there is very minimal tweet activity that can be categorized to spread fake news with a negative context, such as finding supporters for fake news from tweets that are applied.

IV. References

Bralianti, P. D. and Akbar, F. N. (2021) 'Covid-19 Vaccines and its Adverse Events Following Immunization(AEFI) In Indonesia| bralianti | The Avicenna Medical Journal', *The Avicenna Medical Journal*, pp. 19–27.

Chew, C. and Eysenbach, G. (2010) 'Pandemics in the age of Twitter: Content analysis of tweets during the 2009 H1N1 outbreak', *PLoS ONE*, 5(11), pp. 1–13. doi: 10.1371/journal.pone.0014118.

Chhetri, J. K. *et al.* (2020) 'Prevention of COVID-19 in Older Adults: A Brief Guidance from the International Association for Gerontology and Geriatrics (IAGG) Asia/Oceania Region', *Journal of Nutrition, Health and Aging*, 24(5), pp. 471–472. doi: 10.1007/s12603-020-1359-7.

Chou, W. S. and Budenz, A. (2020) 'Considering Emotion in COVID-19 Vaccine Communication : Addressing Vaccine Hesitancy and Fostering Vaccine Confidence Considering Emotion in COVID-19 Vaccine Communication : Addressing Vaccine', *Health Communication*. Routledge, 00(00), pp. 1–5. doi: 10.1080/10410236.2020.1838096.

Deden Ade Nurdeni, Indra Budi, A. B. S. (2021) 'Sentiment Analysis on Covid19 Vaccines in Indonesia : From The Perspective of Sinovac and Pfizer', pp. 122–127.

Du, J. *et al.* (2017) 'Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data', 17(Suppl 2). doi: 10.1186/s12911-017-0469-6.

Featherstone, J. D. *et al.* (2020) 'Exploring childhood anti-vaccine and pro-vaccine communities on twitter – a perspective from influential users', *Online Social Networks and Media*. Elsevier B.V., 20(May), p. 100105. doi: 10.1016/j.osnem.2020.100105.

Gao, Y. *et al.* (2020) 'Inequality in measles vaccination coverage in the "big six" countries of the WHO South-East Asia region', *Human Vaccines and Immunotherapeutics*. Taylor & Francis, 16(7), pp. 1485–1497. doi: 10.1080/21645515.2020.1736450.

Haque, A. and Pant, A. B. (2020) 'E ff orts at COVID-19 Vaccine Development : Challenges and Successes', (December), pp. 1–16.

Hasanah, N. A., Suciati, N. and Purwitasari, D. (2021) 'Identifying Degree-of-Concern on Covid-19 Topics with Text Classification of Twitters', *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, 7(1), pp. 50–62. doi: 10.26594/register.v7i1.2234.

Lucía, A. *et al.* (2018) 'Polarization of the vaccination debate on Facebook', 36, pp. 3606–3612. doi: 10.1016/j.vaccine.2018.05.040.

Miharja, M. *et al.* (2021) 'Implementation of Emergency Public Activity Restrictions (PPKM) in Accordance With Human Rights and Pancasila Principles', *BIRCI-Journal*, (15), pp. 6855–6866.

Nasir, N. M. *et al.* (2021) 'Kebijakan Vaksinasi COVID-19 : Pendekatan Pemodelan Matematika Dinamis Pada Efektivitas Dan Dampak Vaksin Di Indonesia', *Jurnal ABDIMAS*, 4(2), pp. 191–204.

Ophinni, Y. *et al.* (2020) 'COVID-19 Vaccines: Current Status and Implication for Use in Indonesia', *Acta medica Indonesiana*, 52(4), pp. 388–412.

Pulido, C. M. *et al.* (2020) 'A new application of social impact in social media for overcoming fake news in health', *International Journal of Environmental Research and Public Health*, 17(7). doi: 10.3390/ijerph17072430.

Saraswati, I., Muzdalifah, A. and Herawati, A. R. (2021) 'Implementation of Restrictiuons on Community Activities (Ppkm) Policy Analysis Level 1-4 in Dealing With the Covid-19 Outbreak in Indonesia', *International Journal of Social Service and Research*, 1(3), pp. 203–210.

Sinnenberg, L. *et al.* (2016) 'Twitter as a Tool for Health Research : a Systematic Review', pp. 1–8. doi: 10.2105/AJPH.2016.303512.

Subramanian, S. V. and Kumar, A. (2021) 'Increases in COVID-19 are unrelated to levels of vaccination across 68 countries and 2947 counties in the United States', *European Journal of Epidemiology*. Springer Netherlands, (0123456789), pp. 1–4. doi: 10.1007/s10654-021-00808-7.

Supriatna, E. (2020) 'Socio-Economic Impacts of the COVID-19 Pandemic: The Case of Bandung City', *Journal of Governance*, 5(1), pp. 61–70. doi: 10.31506/jog.v5i1.8041.

Tavoschi, L. *et al.* (2020) 'Twitter as a sentinel tool to monitor public opinion on vaccination: an opinion mining analysis from September 2016 to August 2017 in Italy', *Human Vaccines and Immunotherapeutics*. Taylor & Francis, 16(5), pp. 1062–1069. doi: 10.1080/21645515.2020.1714311.

Tfi, M. R., Hamblin, M. R. and Rezaei, N. (2020) 'COVID-19: Transmission, prevention, and potential therapeutic opportunities', *Clinica Chimica Acta*, 508(January), pp. 254–266.

Velavan, T. P. and Meyer, C. G. (2020) 'The COVID-19 epidemic', *Tropical Medicine and International Health*, 25(3), pp. 278–280. doi: 10.1111/tmi.13383.

Veruswati, M. *et al.* (2020) 'Correlation between local eid-al-fitr homecoming (Mudik) with coronavirus

disease-19 during ramadhan season amidst large-scale social distancing in indonesia', *Open Access Macedonian Journal of Medical Sciences*, 8(T1), pp. 570–573. doi: 10.3889/OAMJMS.2020.5369.

Ward, J. K. *et al.* (2020) 'The French public's attitudes to a future COVID-19 vaccine: The politicization of a public health issue', *Social Science and Medicine*, 265, pp. 2016–2021. doi: 10.1016/j.socscimed.2020.113414.