



Classification Algorithm for the Sentiment Analysis of the Covid 19 Pandemic in Indonesia Context

Muhammad Shollahuddin Hamidulloh^{1*}, Abd Kadir Mahamad², and Sharifah Saon²

¹Independent Researcher, 61381 Jakarta, INDONESIA

²Faculty of Electrical and Electronic Engineering,
Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, MALAYSIA

*Corresponding Author

Abstract:

The Covid 19 pandemic that is happening in Indonesia has had a significant impact both in terms of economy, education, and health. Pros and cons related to the policies made by the government in tackling the impact of the spread of the virus produces a variety of sentiments in the community that is interesting to analyze. By taking sentiment data on Twitter, which extracted using Tweepy, then the tweets data are analyzed with Textblob library contained in Python. Naïve Bayes algorithm then, have been used to classify the tweets either in positive, negative, or neutral sentiments. From the experiments conducted, the test results found that there were 60% positive sentiments, 15% negative sentiments and 25% neutral sentiments on economic topics. On the other hand, 70% positive sentiment, 3% negative sentiment and 27% neutral sentiment on the topic of health. Meanwhile, 10% positive sentiment, 50% negative sentiment and 40% neutral sentiment on the topic of education.

Keywords:

Covid 19 · Twitter · Naïve Bayes algorithm · Sentiment analysis · Textblob

1. Introduction

This since the Covid 19 pandemic hit China at the end of 2019. Now the virus has spread throughout the world, not least in Indonesia. The Government of Indonesia established a policy of large-scale social restrictions (PSBB) in April 2020. As a result, offices, schools, restaurants, entertainment venues, markets and other public places were also forced to close, to reduce the process of spreading the virus in the community.

The policy certainly also has a large impact on society. Because of the order to limit social interaction, almost all elements of society feel a significant impact. Like the economic crisis due to the non-rotating economic wheels, academics who require them to do the learning process online, workers are also required to do work from home. On the other hand, limiting social interactions also has an impact on the mental condition of the community, given the fondness of Indonesians to interact and mingle with others.

Pros and cons of the large-scale social restriction policy implemented by the government occur in the discussion space in the community. They polarized between supporting the policies carried out by the government in conducting interaction bans to support the process of handling virus to quickly subside. Or those who reject restrictions on social interaction because it is not accompanied by other policies that support the economy of the middle and lower classes that are most affected in terms of the economy because the economy is not spinning.

Debates and conflicts related to government policies create sentiment in the community. In this case, social media becomes a platform where everyone can provide comments and opinions regarding the government's policy in handling the Covid 19 pandemic [1]. And one of them is Twitter [2-4], which has millions of users in Indonesia who are active every day giving a tweet in the form of information that is factual or just an opinion, making Twitter is a very interesting platform to be used as an object of research related to public sentiment in Indonesia about the policies made by the government to overcoming the Covid 19 pandemic which is currently being hit.

Problems related to the determination of various policies undertaken by the Indonesian government in tackling the Covid 19 pandemic, of course, cause pros and cons in the community. They are polarized, both those who agree and those who disagree with each other to give opinions and information which are then used as a reference in opinion in social media [5].

Various sentiments that emerge in the community in social media related to debates about government policies become interesting things to do analysis. A process of sentiment analysis obtained from data on social media, then classified into three categories: positive, negative or neutral sentiment. From these results can be used as knowledge that can be used as a reference for the government to support decisions to be more precise and strategic when making decisions about the handling of the Covid 19 pandemic. From the analysis of these sentiments will clearly illustrate how to respond to every policy created by the government.

This project focuses on the development of projects that have been used to analyze sentiments on social media (Twitter), especially sentiments towards Indonesian economic, health and education topics related to the problem when the Covid 19 pandemic occurred, using Tweepy to extract data on Twitter, as maximum up to 100,000 tweets in English then analyzed with the TextBlob library before it been classified with the Naïve Bayes algorithm approach to get accurate results.

Sentiment Analysis is one of the techniques in extracting and obtaining information in the form of a person's views (sentiments) on an issue or event. Quoting from the Big Indonesian Dictionary (KBBI), sentiments are "opinions or views based on excessive feelings towards something (contrary to the consideration of the mind)". Sentiment Analysis can be used to reveal public opinion on an issue, service satisfaction, policy, cyber bullying, predicting stock prices, and competitor analysis based on textual data. Every human being generally tends to express response related to an event / situation based on their experience which is the cumulative result of the thought process of the input that has been received by him so far [12]. These responses can generally be categorized into three broad groups, namely positive, negative or neutral responses. However, the phenomenon of exponential data growth is a new challenge in the sentiment analysis process [2][3].

Conventional approaches are no longer the right answer to uncover and determine the type of sentiment in textual data. Employing humans to classify the types of sentiments from a very large and diverse set of textual data certainly require considerable time and money. Imagine if we have millions of tweets per day that must be determined one by one the type of sentiment, certainly not only takes a long time, but also requires enormous resources. Therefore, it need a new technique in sentiment analysis that can automatically extract information from data quickly and responsibly [1].

1.1 Rest API

REST (REpresentational State Transfer) is a standard of web-based communication architecture that is often applied in the development of web-based services. HTTP (Hypertext Transfer Protocol) which is used as a protocol for data communication. REST was first introduced by Roy Fielding in 2000.

In a REST architecture, the REST server has a resource in the form of data and the REST client aims to access and display the resources that will be used by that user. Each resource is identified by URIs (Universal Resource Identifiers) or global IDs. These resources are then represented in various forms such as text format, JSON or XML. In general, the formats that are often used are JSON and XML [4]. Whereas API is an abbreviation of Application Programming Interface. Is a "link" that allows an application to interact with other applications and share data.

1.2 Tweepy

Tweepy is one of the easy-to-use Python libraries to access the API from Twitter. This Tweepy library serves to make it easier for us to get data on Twitter or more simply get Tweet data from Twitter users based on the keywords have been used. So for example if want to get a Tweet about a certain topic, it make that topic into a keyword and then a Tweet that contains the keywords that fill will appear.

Tweepy supports accessing Twitter through Basic Authentication and a newer method, OAuth. Twitter has stopped accepting Basic Authentication so OAuth is now the only way to use the Twitter API. Tweepy provides access to the well-documented Twitter API. With Tweepy, it is possible to get any object and use any method offered by the official Twitter API. For example, a User object has its documentation at <https://dev.twitter.com/docs/platform-objects/users> and following the guidelines, Tweepy can get the appropriate information.

In short, Tweepy is an open source library that provides access to the Twitter API for Python. Although the documentation for Tweepy is rather scarce and doesn't have many examples, the fact that it relies heavily on the Twitter API, which has very good documentation, makes it probably the best Twitter library for Python, especially when considering Streaming API support, which is where Tweepy excels. Other libraries such as Python-Twitter also provide many functions, but Tweepy has the most active and most committed community of code in the last year [13].

1.3 TextBlob

TextBlob is one of the tools or rather the library for processing in the field of "Natural Language Processing (NLP)" using the Python language. Through TextBlob, it can carry out various processes of text data ranging from simple ones such as tokenization (word cutting) to sentiment analysis. Unfortunately TextBlob is not yet implemented for Indonesian so sentiment analysis in Indonesian cannot be done directly. TextBlob is an excellent library for NLP.

TextBlob is actually built on NLTK and Pattern. TextBlob also provides several additional functions than NLTK. This really depends on the type of text analysis you want to do on the data to be processed. But as a general consensus, TextBlob is easier to understand to start with NLP mainly for two reasons - it has very good documentation and interface [5].

TextBlob has good performance when used in text processing, which unstructured data or sentences are being converted into structured form to perform NLP in terms of handling end-to-end interactions. Relationship extraction, semantic decomposition, sentiment analysis, noun phrase extraction are some examples of NLU which are themselves a subset of NLP. Now to work in this area, TextBlob plays a big role that is not very efficient for NLTK.

2. Methodology

This part explains about the rare steps and procedures when developing a project of analysis of the Indonesian people's sentiment towards the Covid 19 pandemic that is being hit as in Fig. 1. The data from Twitter during the lockdown period have been analyzed the sentiment, before in been grouped into positive, negative, or neutral sentiments.

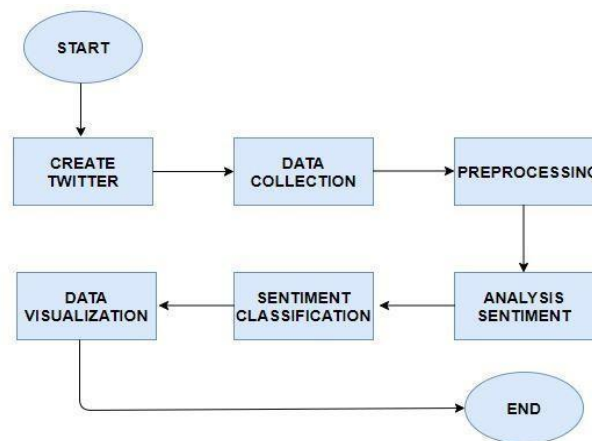


Figure 1: Research flowchart

2.1 Create Twitter

The first thing that needs to be done to build a project sentiment analysis is to register the application to be able to use the Twitter API as in Fig. 2. Application registration can be done by visiting the link <https://developer.twitter.com/>. Select Create New App, to enter a page that contains the form data to fill in. This form contains questions from Twitter about our reasons and goals for becoming a Twitter application developer. The answers sent have been taken into consideration allowing us to take data from Twitter.

Figure 2: Create Twitter

2.2 Data Collection

After Twitter gives access rights in the form of a secret token that can be used as access to retrieve data in Twitter, the next step is to retrieve the data needed for the sentiment analysis process later.

Namely by entering keywords in the form of topics to be analyzed, determining the number of tweets taken and the language used, then the data is displayed with data frames [3][7].

2.3 Preprocessing

Data that has been obtained from the previous crawling process must be cleaned before the sentiment analysis process is carried out. Data cleaning process is carried out with the aim to eliminate characters, symbols and letters that can affect the analysis process [3]. Data sharing also deletes Twitter users' retweets to get an original and more meaningful tweet for the sentiment analysis process [8].

2.4 Sentiment Analysis Process

This sentiment analysis process is carried out using TextBlob to calculate and find polarity and subjectivity as in Fig. 3 and 4. Sentiment analysis step is done by adding subjectivity and polarity of tweets to DataFrame. To do this, the step taken is to create two functions: one to get a tweet called Subjectivity (how subjective or argumentative the text - a score of 0 is a fact, and a score of +1 is very much an opinion) and the other is to get a tweet called Polarity (how positive or negative text, which score -1 is the highest negative score, and +1 score is the highest positive score) [9]. For each word, it get a sentiment score (how positive / negative) and subjectivity score (how much they think).

Polarity is how positive or negative a word is. (-1) very negative. Or (+1) very positive. While, subjectivity is how subjective, or argues a word. Where (0) is fact, and (+1) there are very many opinions.

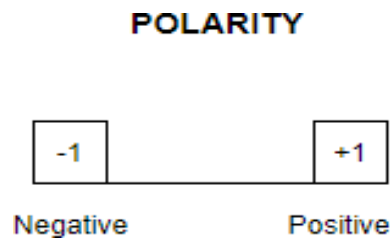


Figure 3: Polarity

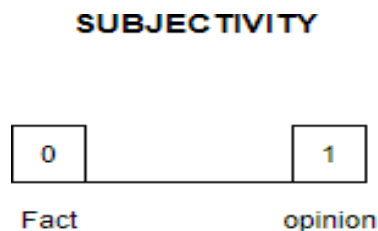


Figure 4: Subjectivity

In this step, the process of extracting Twitter data is done by using Tweepy. Sentiment analysis of tweets that have been extracted have been processed using the TextBlob library with the Python programming language.

2.5 Sentiment Classification

The process of classifying sentiments that have been analyzed aims to determine or categorize a sentiment as good or bad, like or dislike, agree or disagree. This process is done by creating a function to calculate the negative analysis (-1), neutral (0), and positive (+1), and add information to a new column called Analysis. Then, display the results in the data frame [10][11][14].

2.6 Data Visualization

The data that has been obtained from the sentiment analysis process is of course still in the form of text and numbers in the data frame. This becomes difficult to understand if it is not converted into a more attractive visual form. The data visualization process aims to make it easier to draw conclusions from data that has been processed and analyzed. The data have been visualized in the form of chart, diagram or table as shown in Fig. 5. This data visualization process uses the Matplotlib library which is already available in Python [15].



Figure 5: Data visualization

3. Result and Analysis

This part explains about the implementation of sentiment analysis on Covid 19, which takes Indonesian sentiment data on Twitter using Tweepy which have been analyzed with Textblob as described in the previous chapter. In this chapter three topics related to sentiment towards Covid 19 have been consider, namely the sentiment based on the education, health and economics. The process has been explained in sequence starting from the process of making the Twitter application to the analysis process.

3.1 Crawling

The first step that needs to be done to get the data acquisition for the analysis process is to register Twitter developer account. This account is useful to get access rights in the form of a secret token to retrieve the data that is required as shown in Fig. 6.

```
consumer_key= 'oU922cUF1hOn1JKRor45fKAhx'
consumer_secret= 'ubVmAdjLsEmrhS1PDyRfeL46Y1hVe36103syn3YSDXNQvOwMtr'

access_token='1228095171847323648-YtpheYWjasBdkFmLohhMUP2V21UHsT'
access_token_secret='n1VQAdFoomuRL3aVnkrISGUM6Lh9NDVIFrbpy4eRyYdLF'

auth = tweepy.OAuthHandler("oU922cUF1hOn1JKRor45fKAhx", "ubVmAdjLsEmrhS1PDyRfeL46Y1hVe36103syn3YSDXNQvOwMtr")
auth.set_access_token("1228095171847323648-YtpheYWjasBdkFmLohhMUP2V21UHsT", "n1VQAdFoomuRL3aVnkrISGUM6Lh9NDVIFrbpy4eRyYdLF")

api = tweepy.API(auth)
```

Figure 6: Access token secret

After getting the access rights in the form of tokens, the next step is to determine the keywords that match the tweets as needed, also enter the number of tweets that have been taken later. Fig. 7 shows the data that has been obtained is displayed into a data frame for easier observation.

```
# Extract tweets from the twitter user
posts = api.search(q=['health+covid+indonesia'], count= "100000", lang ="en")

print("10 recent tweets:\n")
for tweets in posts[:10]:
    print(tweets.text + '\n')
```

Out[38]:

	Tweets
0	RT @WHOIndonesia: #WHOIndonesia has been suppo...
1	RT @WHOIndonesia: #WHOIndonesia has been suppo...
2	RT @WHOIndonesia: #WHOIndonesia has been suppo...
3	RT @WHOIndonesia: #WHOIndonesia has been suppo...
4	RT @WHOIndonesia: #WHOIndonesia has been suppo...
5	RT @WHOIndonesia: #WHOIndonesia has been suppo...
6	RT @WHOIndonesia: #WHOIndonesia has been suppo...
7	RT @PhuketDailyNews: Phuket Covid-19: Indonesi...
8	RT @PhuketDailyNews: Phuket Covid-19: Indonesi...
9	RT @PhuketDailyNews: Phuket Covid-19: Indonesi...

Figure 7: Data frame extraction

3.2 Processing

The data that has been obtained from the process of extracting data on Twitter, is carried out through the data cleaning process. This process aims to eliminate the symbols and characters in a tweet that may affect the results of the analysis later so that it runs well and gets the desired results [9]. Fig. 8 shows frame data that has gone through the data cleaning process.

The next step is to calculate the score of each tweet that has been obtained. Sentiment analysis is done with the Textblob library in Python, determining the subjectivity and polarity of a tweets by looking at each word, to get a polarity score (how positive / negative) and subjectivity score (how much their opinion) as shown in Fig. 9.

```
In [39]: # Create a function to clean the tweets
def cleanTxt(text):
    text = re.sub('@[A-Za-z0-9]+', '', text) #Removing @mentions
    text = re.sub('#', '', text) # Removing '#' hash tag
    text = re.sub('RT[\s]+', '', text) # Removing RT
    text = re.sub('https?:\/\/\S+', '', text) # Removing hyperlink

    return text

# Clean the tweets
df['Tweets'] = df['Tweets'].apply(cleanTxt)

# Show the cleaned tweets
df.head(10)
```

Out[39]:

	Tweets
0	: WHOIndonesia has been supporting and to tr...
1	: WHOIndonesia has been supporting and to tr...
2	: WHOIndonesia has been supporting and to tr...
3	: WHOIndonesia has been supporting and to tr...
4	: WHOIndonesia has been supporting and to tr...
5	: WHOIndonesia has been supporting and to tr...
6	: WHOIndonesia has been supporting and to tr...
7	: Phuket Covid-19: Indonesia posts highest dai...
8	: Phuket Covid-19: Indonesia posts highest dai...
9	: Phuket Covid-19: Indonesia posts highest dai...

Figure 8: Frame data that has gone through the data cleaning process

```
In [40]: # Create a function to get the subjectivity
def getSubjectivity(text):
    return TextBlob(text).sentiment.subjectivity

# Create a function to get the polarity
def getPolarity(text):
    return TextBlob(text).sentiment.polarity

# Create two new columns 'Subjectivity' & 'Polarity'
df['Subjectivity'] = df['Tweets'].apply(getSubjectivity)
df['Polarity'] = df['Tweets'].apply(getPolarity)

# Show the new dataframe with columns 'Subjectivity' & 'Polarity'
df.head(10)
```

Figure 9: Subjectivity & polarity

In Fig. 10 shows score results from the calculation of subjectivity and polarity of the tweets obtained.

Out[40]:

	Tweets	Subjectivity	Polarity
0	: WHOIndonesia has been supporting and to tr...	0.375	0.375
1	: WHOIndonesia has been supporting and to tr...	0.375	0.375
2	: WHOIndonesia has been supporting and to tr...	0.375	0.375
3	: WHOIndonesia has been supporting and to tr...	0.375	0.375
4	: WHOIndonesia has been supporting and to tr...	0.375	0.375
5	: WHOIndonesia has been supporting and to tr...	0.375	0.375
6	: WHOIndonesia has been supporting and to tr...	0.375	0.375
7	: Phuket Covid-19: Indonesia posts highest dai...	0.000	0.000
8	: Phuket Covid-19: Indonesia posts highest dai...	0.000	0.000
9	: Phuket Covid-19: Indonesia posts highest dai...	0.000	0.000

Figure 10: Display subjectivity & polarity

The next analysis process is done by classifying the scores from the polarity calculation into 3 categories, namely tweets are neutral if the score is equal to 0, positive if the score is more than 0 and negative if the score is less than 0 as shown in Fig. 11.


```
In [41]: # Create a function to compute negative (-1), neutral (0) and positive (+1) analysis
def getAnalysis(score):
    if score < 0:
        return 'Negative'
    elif score == 0:
        return 'Neutral'
    else:
        return 'Positive'
df['Analysis'] = df['Polarity'].apply(getAnalysis)
# Show the dataframe
df.head(10)
```

Figure 11: Analysis score

In Fig.12 shows the results of the process of categorizing analysis sentiments whether positive or neutral.

```
Out[41]:
```

	Tweets	Subjectivity	Polarity	Analysis
0	: WHOIndonesia has been supporting and to tr...	0.375	0.375	Positive
1	: WHOIndonesia has been supporting and to tr...	0.375	0.375	Positive
2	: WHOIndonesia has been supporting and to tr...	0.375	0.375	Positive
3	: WHOIndonesia has been supporting and to tr...	0.375	0.375	Positive
4	: WHOIndonesia has been supporting and to tr...	0.375	0.375	Positive
5	: WHOIndonesia has been supporting and to tr...	0.375	0.375	Positive
6	: WHOIndonesia has been supporting and to tr...	0.375	0.375	Positive
7	: Phuket Covid-19: Indonesia posts highest dai...	0.000	0.000	Neutral
8	: Phuket Covid-19: Indonesia posts highest dai...	0.000	0.000	Neutral
9	: Phuket Covid-19: Indonesia posts highest dai...	0.000	0.000	Neutral

Figure 12: Display analysis

3.3 Processing

Fig. 13 shows the results of the sentiment analysis process that have been carried out are then displayed on a chart to facilitate the process of observing a sentiment. Meanwhile, Fig. 14 shows the process for displaying the calculation value of the analyzed data. The data is then visualized into a bar chart to facilitate the process of observing the results of sentiment analysis as in Fig. 15.

```
In [6]: # Let's plot the results
import matplotlib.pyplot as plt
%matplotlib inline
plt.rcParams['figure.figsize'] = [10, 8]
for i in range(0, df.shape[0]):
    plt.scatter(df["Polarity"][i], df["Subjectivity"][i], color='Blue')

plt.title('Sentiment Analysis', fontsize = 20)
plt.xlabel('← Negative - - - - - Positive →', fontsize=15)
plt.ylabel('← Facts - - - - - Opinions →', fontsize=15)
plt.show()
```

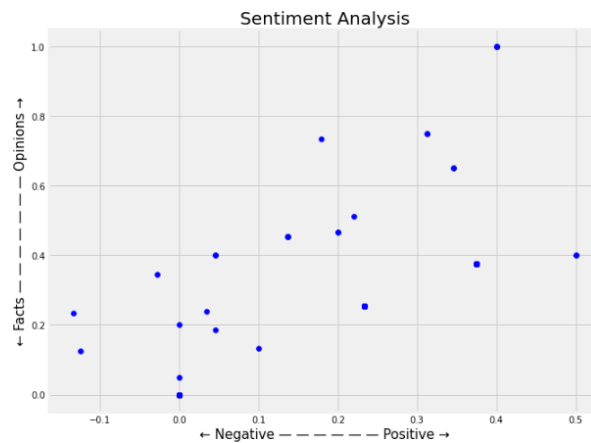


Figure 13: Scatter chart

```
In [11]: # Show the value counts
df['Analysis'].value_counts()
```

```
Out[11]: Positive    70
         Neutral     27
         Negative     3
         Name: Analysis, dtype: int64
```

Figure 14: Count value

```
In [12]: # Plotting and visualizing the counts
plt.title('Sentiment Analysis')
plt.xlabel('Sentiment')
plt.ylabel('Counts')
df['Analysis'].value_counts().plot(kind = 'bar')
plt.show()
```

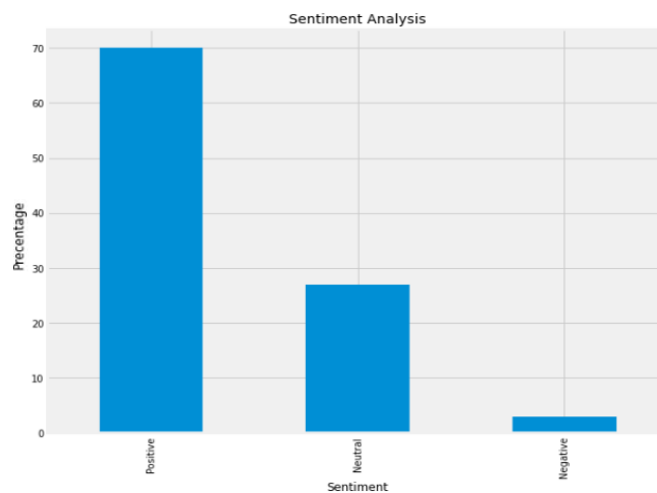


Figure 15: Bar chart

3.4 Result Analysis

The process of sentiment analysis about the sentiments of the Indonesian people in relation to the current Covid 19 pandemic, taking data sourced from Twitter. The data is extracted using Tweepy and then analyzed with the Textblob library in Python. By using the Naïve Bayes approach to classify the results of the analysis into 3 categories, namely positive, negative or neutral.

In the discussion of this process, 3 topics have been developed in the community at the time of the current Covid 19 pandemic, namely community sentiment about economic conditions, health and education. By taking the amount of data as many as 100,000 tweets in each category. It's just taken tweets using English because the process of analysis using Textblob can only process data with English and French only.

3.4.1 Economic Sentiment

The first sentiment that have been analyzed is the community's sentiment related to economic conditions that are happening in the in the middle of the Covid 19 pandemic. The process is done by entering the words 'economy' as a keyword to look for tweets related to the economy, then the word 'covid' as a keyword to search for tweets related to the Covid 19 virus condition and 'Indonesia' as a keyword to search for tweets that are relating to the conditions of the country being analyzed as shown in Fig. 16 and Fig. 17.

```
In [51]:
1 # Show the value counts
2 df['Analysis'].value_counts()
```

```
Out[51]: Positive    60
         Neutral    25
         Negative   15
         Name: Analysis, dtype: int64
```

Figure 16: Count value economic sentiment

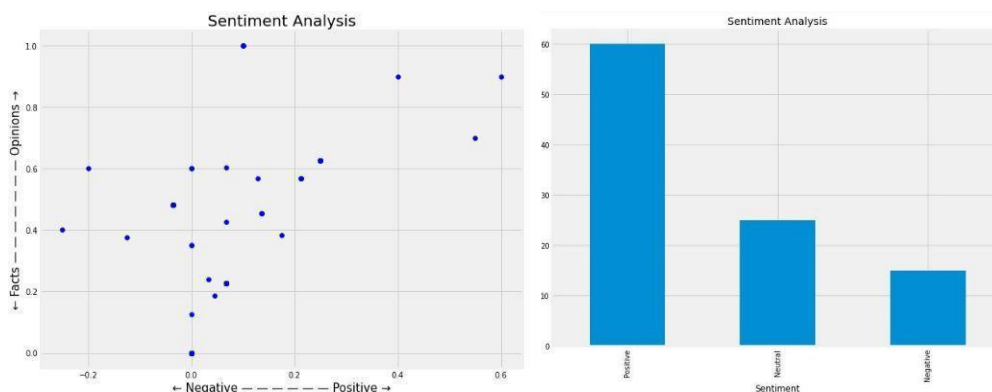


Figure 17: Chart economic sentiment

3.4.2 Health Sentiment

The second sentiment that have been analyzed is the community's sentiment related to the current health condition in the midst of the Covid 19 pandemic. The process is done by entering the word 'health' as a keyword to look for tweets related to health, then the word 'covid' as a keyword to search for tweets related to the Covid 19 virus condition and 'Indonesia' as a keyword to search for tweets that are relating to the conditions of the country being analyzed.

From the sentiment analysis process that was carried out on the topic of the health the following results were obtained: there were 70% positive sentiments, 27% neutral sentiment tweets and 3% tweets with negative sentiments as in Fig. 18 and Fig. 19.

```
In [11]: # Show the value counts
df['Analysis'].value_counts()

Out[11]: Positive    70
         Neutral     27
         Negative     3
         Name: Analysis, dtype: int64
```

Figure 18: Count value health sentiment

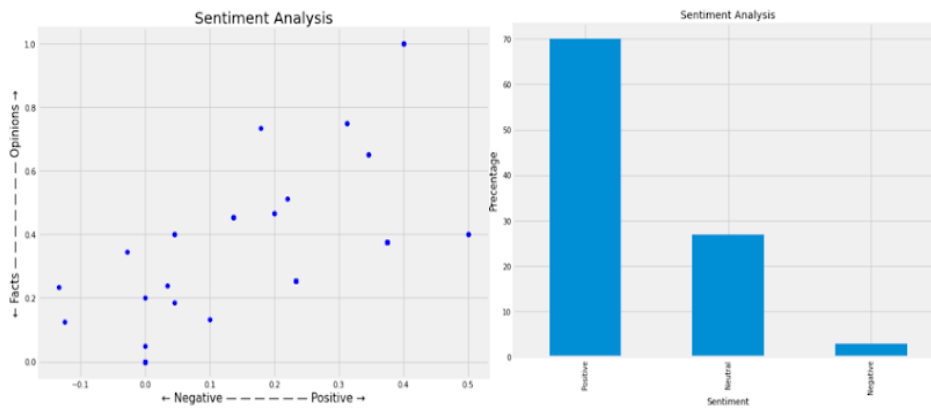


Figure 19: Chart health sentiment

3.4.3 Education Sentiment

The third sentiment to be analyzed is the community's sentiment regarding the current state of education in the midst of the Covid 19 pandemic. The process is done by entering the words 'education' as a keyword to look for tweets related to education, then the word 'covid' as a keyword to search for tweets related to the condition of the Covid 19 virus and 'Indonesia' as a keyword to search for tweets that are relating to the conditions of the country being analyzed.

From the sentiment analysis process that was carried out on the topic of the education the following results were obtained: there were 50% negative sentiments, 40% neutral sentiment tweets and 10% tweets with positive sentiments as shown in Fig. 20 and Fig. 21.

```
In [13]: # Show the value counts
df['Analysis'].value_counts()

Out[13]: Negative    5
         Neutral     4
         Positive     1
         Name: Analysis, dtype: int64
```

Figure 20: Count value education sentiment

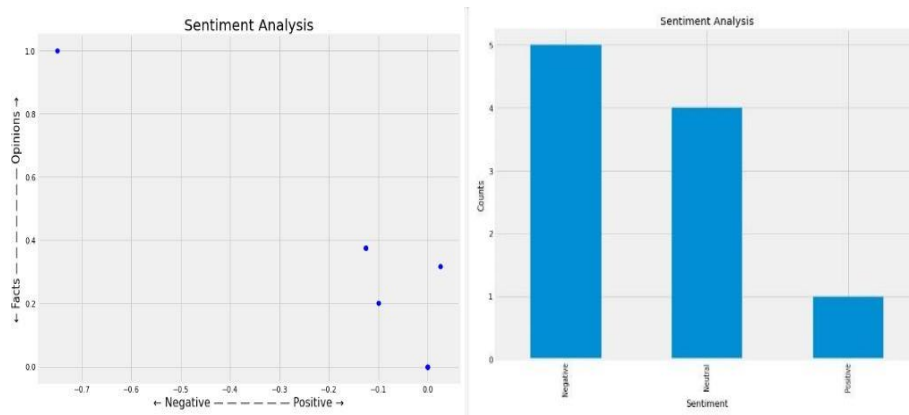


Figure 21: Chart education sentiment

4. Conclusion

This project was made with the aim of obtaining the results of a community sentiment analysis process in Indonesia regarding the Covid 19 pandemic that is currently underway. By taking the existing sentiment data on Twitter with a TextBlob library in Python. With the Naïve Bayes algorithm approach to classify the results of the analysis into sentiments with positive, negative or neutral categories. From the results of the analysis that has been done, the results for each topic are 60% positive, 15% negative and 25% neutral for economic sentiment. 70% positive, 27% negative and 3% neutral for health sentiment and 10% positive, 50% negative and 10% neutral for education sentiment. The conclusions that have been obtained from Indonesian sentiments related to the 3 sentiment topics that have been tested are more likely to be positive, which this can be seen in the results of sentiment analysis on economic and educational topics, which is 60% and 70% positive sentiments. while negative sentiment is only on the topic of education, which has the largest negative sentiment at 50%. while the rest is on sentiment towards the economy, education and health, there are also many neutral sentiments. For future work the sentiment analysis will consider tweet in Indonesia language rather than English.

References

- [1] Samuel, J., Ali, G.G.N., Rahman, M., Esawi, E., Samuel, Y.: COVID 19 Public Sentiment Insights and Machine Learning for Tweets Classification. *Information*, 11(6), 314, (2022). doi: 10.3390/info11060314.
- [2] Alsaedi, A., Zubair Khan, M.: A Study on Sentiment Analysis Techniques of Twitter Data. *International Journal of Advanced Computer Science and Applications*, 10 (2), pp. 361-374 (2019). doi: 10.14569/IJACSA.2019.0100248.
- [3] Zulfa, I., Winarko, E.: Sentimen Analisis Tweet Berbahasa Indonesia dengan Deep Belief Network. *Indonesian Journal of Computing and Cybernetics Systems*. 11(2), (2017). <https://doi.org/10.22146/ijccs.24716>
- [4] Ratnawati, F.: Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter. *Jurnal INOVTEK Polbeng*. 3(1), (2018). <https://doi.org/10.35314/isi.v3i1.335>
- [5] Kunal, S., Saha, A., Varma, A., Tiwari, V.: Textual Dissection Of Live Twitter Reviews Using Naive Bayes. *Procedia Comput. Sci.*, 132, pp. 307–313 (2018). doi: 10.1016/j.procs.2018.05.182.
- [6] Arianto, D., Pui, N.: Social Media Analysis: Utilization of Social Media Data for research on COVID 19. April 2020.

- [7] Arief, R., Imanuel, K.: ANALISIS SENTIMEN TOPIK VIRAL DESA PENARI PADA MEDIA SOSIAL TWITTER DENGAN METODE LEXICON BASED. *Jurnal Ilmiah Matrik*, vol. 21, no. 3, pp. 242–250 (2019). <https://doi.org/10.33557/jurnalmatrik.v21i3.727>
- [8] Nafan, M.Z., Amalia, A.E.: Kecenderungan Tanggapan Masyarakat terhadap Ekonomi Indonesia berbasis Lexicon Based Sentiment Analysis. *Jurnal Media Informatika Budidarma*. vol. 3, no. 4, pp. 268– 273 (2019). doi: 10.30865/mib.v3i4.1283.
- [9] Sinaga, A.: Utilization of Sentiment Analysis for Tourist. *Jurnal Penelitian Pos dan Informatika Pemingkatan Popularitas Tujuan Wisata*. vol. 7, no. 2, pp. 109–120 (2017). doi: 10.17933/jppi.2017.070203.
- [10] Khaira, U., Johanda, R., Eko, P., Utomo, P., Suratno, T.: Sentiment Analysis of Cyberbullying on Twitter Using SentiStrength. *Indonesian Journal of Artificial Intelligence and Data Mining*. vol. 3, no. 1, pp. 21–27 (2020). Doi:10.24014/ijaidm.v3i1.9145
- [11] Kurniawan, T.: Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Media Mainstream Menggunakan Naïve Bayes Classifier Dan Support Vector Machine. Undergraduate Thesis (2017).
- [12] Habibi, R., Setyohadi, D.B., Wati, E.: Analisis Sentimen Pada Twitter Mahasiswa Menggunakan Metode Backpropagation. *J. Inform.*, vol. 12, no. 1, pp. 103–109 (2016).
- [13] Monarizqa, N., Nugroho, L.E., Hantono, B.S.: Penerapan Analisis Sentimen Pada Twitter Berbahasa Indonesia Sebagai Pemberi Rating. *J. Penelit. Tek. Elektro dan Teknol. Inf.*, vol. 1, pp. 151–155 (2014).
- [14] Cahyono, Y.: Analisis Sentiment pada Sosial Media Twitter Menggunakan Naïve Bayes Classifier dengan Feature Selection Particle Swarm Optimization dan Term Frequency. *J. Inform. Univ. Pamulang*, vol. 2, no. 1, p. 14 (2017).
- [15] Kabir M.Y., Madria, S.: CoronaVis: A Real-time COVID 19 Tweets Analyzer and Data Repository. *Social and Information Networks*. 2020. <https://doi.org/10.48550/arXiv.2004.13932>