

# Indonesian Twitter Sentiment Analysis Application on The Covid19 Vaccine Using Naive Bayes Classifier

1<sup>st</sup> Adhithia Erfina  
*Departement of System Infomation*  
Nusa Putra University  
Sukabumi, Indonesia  
[adhithia.erfina@nusaputra.ac.id](mailto:adhithia.erfina@nusaputra.ac.id)

2<sup>nd</sup> Moneyta Dholah Rosita NDK  
*Departement of System Infomation*  
Nusa Putra University  
Sukabumi, Indonesia  
[moneyta.dholah\\_si17@nusaputra.ac.id](mailto:moneyta.dholah_si17@nusaputra.ac.id)

3<sup>rd</sup> Rahmat Hidayat  
*Departement of System Infomation*  
Nusa Putra University  
Sukabumi, Indonesia  
[rahmat.hidayat\\_si17@nusaputra.ac.id](mailto:rahmat.hidayat_si17@nusaputra.ac.id)

4<sup>th</sup> Aris Subagja  
*Departement of System Infomation*  
Nusa Putra University  
Sukabumi, Indonesia  
[aris.subagja\\_si17@nusaputra.ac.id](mailto:aris.subagja_si17@nusaputra.ac.id)

5<sup>th</sup> Haerul Ramadhan  
*Departement of System Infomation*  
Nusa Putra University  
Sukabumi, Indonesia  
[haerul.ramadhan\\_si17@nusaputra.ac.id](mailto:haerul.ramadhan_si17@nusaputra.ac.id)

6<sup>th</sup> CSA Teddy Lesmana  
*Departement of Law*  
Nusa Putra University  
Sukabumi, Indonesia  
[teddy.lesmana@nusaputra.ac.id](mailto:teddy.lesmana@nusaputra.ac.id)

7<sup>th</sup> Sudin Saepudin  
*Departement of System Infomation*  
Nusa Putra University  
Sukabumi, Indonesia  
[sudin.saepudin@nusaputra.ac.id](mailto:sudin.saepudin@nusaputra.ac.id)

8<sup>th</sup> Muhamad Muslih  
*Departement of System Infomation*  
Nusa Putra University  
Sukabumi, Indonesia  
[muhamad.muslih@nusaputra.ac.id](mailto:muhamad.muslih@nusaputra.ac.id)

**Abstract**— It's even one year since the COVID-19 pandemic hit Indonesia, to anticipate it, the government brought in a COVID-19 vaccine. Various types of COVID-19 vaccine have been introduced to Indonesia, including which ones will be considered the best according to the community through the Twitter platform. One of the venues that creates the most public sentiment is Twitter. It can be determined whether the public fully approves or rejects the existence of vaccination in Indonesia by analyzing public sentiment surrounding the COVID-19 vaccine. Data acquisition using a crawling procedure by connecting the Twitter API, pre-processing, sentiment categorization, and sentiment analysis outcomes are the stages of the sentiment analysis process to become a sentiment analysis application. The PHP and MySQL programming languages are used to create the database for the sentiment analysis application. After the application has been fully implemented, it can do sentiment analysis from each dictionary probability using the Naive Bayes Classifier approach. The study of the two keywords "vaksin covid" and "vaksin corona" yielded the following results. It has 93% positive sentiment results, 72% negative sentiment results, and 35% neutral sentiment outcomes, with an accuracy of 94.74% and 75.47% per keyword. Meanwhile, the Sinopharm vaccine, which has the most positive attitude with the terms "vaksin sinovac," "vaksin astrazeneca," "vaksin sinopharm," and "vaksin nusantara," has 84 percent tweets with a 74.23% accuracy rate.

**Keywords**— Vaccine, COVID19, Indonesian Twitter, naïve bayes classifier.

## I. INTRODUCTION

After a year of being hit by the COVID-19 Pandemic, the President of the Republic of Indonesia announced on the Presidential Secretariat's Youtube Channel that he would bring in the COVID-19 Vaccine as a result of his diplomacy

with several countries that have succeeded in developing a COVID-19 vaccine and administering it regularly. free.

Following the immunization, there were many positive and negative reactions from the community, as well as complaints about some positive and negative side effects. Since January 13, 2021, when the national immunization campaign was launched to combat COVID-19, Indonesia has used three types of vaccines: Sinovac (CoronaVac), AstraZeneca manufactured by SK Bio, and Sinopharm[1]. Furthermore, the Indonesian vaccine, dubbed Vaccine Nusantara, will be completed by the end of the year, according to the health ministry.

Which of these vaccines will be deemed the best by the community? An analysis of community sentiment surrounding the COVID-19 vaccine will reveal whether the community fully approves or rejects the existence of a vaccine in Indonesia, as well as which vaccine type will be accepted by the Indonesian people. Because every positive and negative feeling will have an impact on the government and the implementation of this vaccination, ensuring that Indonesians are better protected against COVID-19 in the future.

One of the areas where people discuss the COVID19 vaccine is Twitter. Twitter is one of five popular platforms in Indonesia, along with Youtube, Whatsapp, Facebook, and Instagram. Not only that, but Twitter is a communication tool that the general public may use to express themselves, and it is becoming increasingly popular around the world, including in Indonesia. Every day, the number of tweets from Twitter users grows, and the data that emerges on Twitter can reveal what is going on in society. As a result, there is data from Twitter tweets that may be analyzed for the sentiment.[2]

In addition, determining public opinion on the COVID-19 vaccine needs a system that can categorize each opinion. The Naive Bayes Classifier was chosen by the researcher (NBC). The probabilistic Naive Bayes Classifier-based approach has several advantages over the numerical-based approaches, including being simple, fast, and accurate. For text classification, the Naive Bayes Classifier approach uses word properties that appear in a document as the basis for categorization. Written in Rish (2001), although the assumption of independence between words in a document cannot be entirely met, the Naive Bayes Classifier performs relatively well in classification. So this study will determine the value of positive, negative, or neutral sentiment from a collection of tweets from Twitter users[2].

## II. LITERATUR REVIEW

Written in Yulianton[3] analyzes conversations about venting negative feelings on social media Twitter. To acquire data from Twitter, we'll utilize the R programming language, which offers a package for doing so. The findings of this study are data on e-women Venting Negative Feelings that may be utilized to help internet users make informed selections about which internet service provider to choose.

Testing of web-based implementation using the PHP programming language reveals that tweets may be categorized automatically, according to Ardiani's[4] written. The data is divided into three (three) categories: positive, negative, and neutral. The process of conducting sentiment analysis begins with text processing; after the text processing stage, opinions are classified using the Naive Bayes classification method into positive, negative, or neutral categories. Overall, the test results attained an accuracy of 72 percent by using the Naive Bayes classification approach to automatically categorize sentiment information contained in twitter data into three classes, namely positive, negative, and neutral, with 450 training data and 50 test data.

According to the library that has been supplied and utilizing the PHP tweet programming language, conversations on Twitter can be evaluated and produce certain thoughts or feelings and can be divided automatically into three types, namely positive, negative, and neutral. As a result, both are accurate enough to be employed in the sentiment analysis of the Covid-19 vaccination.

The performance of the Naive Bayes and K-NN algorithms in classifying with the Lexicon dictionary approach is compared in written Azhar[5]. The goal of this research is to compare the Lexicon approach to the Naive Bayes and K-NN algorithms in terms of accuracy, precision, and recall. The combination of the Naive Bayes algorithm and the Lexicon method has a better performance with an accuracy of 81 percent, followed by K-NN with an accuracy of 77 percent, according to the evaluation.

Padmaja[6] contrasts two sentiment analysis approaches, the Naive Bayes Classifier, and the Support Vector Machine, in a written paper. The United Progressive Alliance (UPA), Telugu Desam Party (TDP), and Telangana Rashtra Samithi were the three contesting parties in the 2009 Indian elections, and 689 news articles were collected for this study (TRS). According to the findings of this investigation, the Naive Bayes Classifier calculation appears to be superior, with an

accuracy of 72 percent compared to 70 percent for the Support Vector Machine.

Both studies prove that the Naive Bayes method performs very good classification in sentiment analysis whether using a library or not, it is proven through the high level of accuracy of the comparison methods. So the researchers used the Naive Bayes method in this study.

### Vaccines COVID-19

Vaccines are biological products containing antigens that, when given to a person, will actively develop specific immunity against certain diseases. The COVID-19 vaccine is not a drug. Vaccines encourage the formation of specific immunity in COVID-19 to avoid contracting or possibly becoming seriously ill. As long as a safe and effective vaccine has not been found, the protective measures we can do are 3M discipline: Wear masks properly, keep your distance and stay away from crowds, and Wash your hands with running water and soap [7].

There have been several types of vaccines circulating in Indonesian society until now, as of May 1, 2021, and it has been confirmed that they have been imported, namely the Sinovac Vaccine from China, the Astrazeneca Vaccine from England, the Sinopharm Vaccine from China, and the Indonesian-made Nusantara Vaccine.

### Sentiment Analysis

Sentiment analysis is a study in the field of computing that uses textual representations of opinions, opinions, and emotions [8].

Existing text in sentences, opinions, opinions, and documents can be classified using sentiment analysis. Positive, negative, or neutral feelings can be expressed in this way [9]. Sentiment analysis methods are considered effective for getting information about whether other people have a positive or negative view of a topic[10].

### Twitter and Twitter API

Twitter is a social media platform and online microblogging service that allows users to send and receive text-based messages of up to 140 characters, known as tweets [11].

Developers can create applications based on Twitter's Application Programming Interface (API), which allows them to cater to their demands. The Twitter API documentation can be obtained at <http://dev.twitter.com>. Twitter APIs come in a variety of flavors, including:

1. Twitter REST API: Twitter REST and Twitter Search are both parts of the Twitter REST API. Twitter REST gives you access to fundamental data and Twitter objects. To find an instance of the Twitter object or to seek trends, use the Twitter search methods.
2. Twitter Streaming API: This API is commonly used for extracting data because through this API information can be obtained in real time with a very high volume [13].

### Naive Bayes Classifier

The Naive Bayes Classifier is a probabilistic approach. Although the Naive Bayes Classifier algorithm is simple, it provides good accuracy and performance values when it

comes to categorizing text [14]. The Naive Bayes Classifier can be used to categorize an opinion as favorable, negative, or neutral[15].

The advantage of employing the Nave Bayes Classifier method in the text classification process is that the method requires little training data. The Naive Bayes Classifier equation [11] is found in equation (1).

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Ket :

- X : Data with an unknown class.
- H : The data hypothesis is class specific.
- P(H|X) : Probability H based on condition X.
- P(H) : Probability H.
- P(X|H) : Probability of X based on the conditions of H.
- P(X) : Probability X.

### The Naive Bayes Classification Method

The Naive Bayes Classifier method requires various instructions to determine which class is appropriate for examining a sample throughout the classification process. As a result, equation (1) is changed as a result of equation (2) [11].

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)} \quad (2)$$

The following is the Naive Bayes Classifier theorem model that will be utilized in the classification process:

$$P(F_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i) \quad (3)$$

### Naive Bayes Method Test Results

The system's accuracy in carrying out the classification process is then tested using the results of the classification process. Equation (4) shows how to test the Naive Bayes Classifier algorithm as a result of document classification. This test will be carried out by manually calculating data that already contains labels with the system's results [16].

$$P(P(NA)) = \frac{\text{Data}}{X} \times 100\% \quad (4)$$

Where:

- P(NA) : Percentage of accuracy value
- Data : the number of test data that is true
- X : total sample size

### III. RESEARCH METHODOLOGY

This study uses the Naive Bayes Classifier method in analyzing the opinions or opinions of Twitter users on the COVID-19 vaccine. The flow diagram in this study using the Naive Bayes Classifier method is as follows:

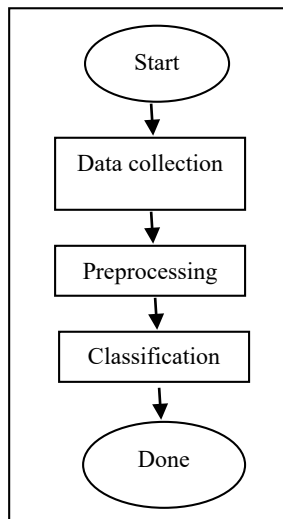


Figure 1 Flow of Sentiment Analysis

#### A. Data Collection

The dataset used in this study is a collection of opinions in the form of tweets about the covid-19 vaccine in the period 01 May 2021 which were taken via Twitter API. All tweet data is loaded into the MySQL database.

#### B. Preprocessing

Before entering the classification process, tweet data goes through the preprocessing stage to transform data from an unstructured form into structured data so that the analysis process becomes easier [17]. The method applied in the preprocessing process can be seen in Figure 2.

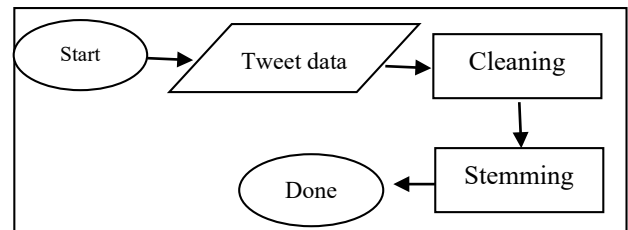


Figure 2. Preprocessing flow

#### • Cleaning

Cleaning in this study is aimed at reducing noise and converting the entire text in the document into lowercase letters [18]. As:

TABLE 1. CLEANING RESULTS

Before Cleaning	After Cleaning
@DalamIstana Bg, bahas donk tentang vaksin corona.... Gw bingung dgn negeri ini... Koq tiba2 baik dgn menggratiskan vaksin corona kerakyat? Sementara dihal lain zalimnya kerakyat luar biasa... Sangat... Sangat mencurigakan ni vaksin sinovac dan astrazaneca	Bg, bahas donk tentang vaksin Gw bingung dgn negeri Koq tiba2 baik dgn menggratiskan vaksin corona kerakyat? Sementara dihal lain zalimnya kerakyat luar Sangat mencurigakan ni vaksin sinovac dan astrazaneca

#### • Stemming

The stemming stage is the process of finding the root word by removing all affixes, both in the form of prefixes, suffixes, and a combination of prefixes and suffixes that exist in each word in the data [18].

TABEL 2. HASIL STEMMING

Before Stemming	After Stemming
Bg, bahas donk tentang vaksin Gw bingung dgn negeri Koq tiba2 baik dgn menggratiskan vaksin corona kerakyat? Sementara dihal lain zalimnya kerakyat luar Sangat mencurigakan ni vaksin sinovac dan astrazeneca	bg bahas donk tentang vaksin gw bingung dgn negeri koq tiba2 baik dgn gratis vaksin corona rakyat sementara hal lain zalim rakyat luar sangat curiga ni vaksin sinovac dan astrazeneca

### C. Classification

After completing data collection and preprocessing, all data will enter the classification stage using the Naïve Bayes Classifier algorithm using the dictionary or library that has been provided. So that the tweet data will be automatically labeled positive, negative or neutral. After that it will be automatically inputted into the database that has been provided.

### D. Results of the Analysis

Measuring the test results of the Naive Bayes algorithm and doing a descriptive analysis of each searched phrase in the form of a bar and pie chart is the final step in this research. So that a clear percentage may be used to compare positive, negative, and neutral sentiments. The Covid19 vaccine case study was utilized to evaluate the application, along with seven keywords: "Vaksin Covid", "Vaksin Corona", "Vaksin Sinovac", "Vaksin AstraZeneca", "Vaksin Sinopharm", and "Vaksin Nusantara". The application's output is a forecast of each tweet's polarity, which is then compared to the findings of each tweet's assessment in the database, which is done manually (hand-labeled).

### E. System Architecture

The system built is a system that can be used in conducting sentiment analysis on public responses to the Covid-19 vaccine on Twitter. The system works by crawling tweets using the Twitter API, which will be processed by the system to get the sentiment values contained in the tweet by going through the text preprocessing stage, then after that the sentiment classification will be carried out to get the classification results.

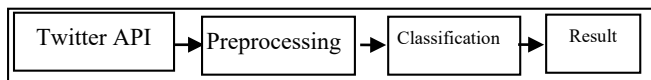


Figure 3. System Architecture

## IV. IMPLEMENTATION AND RESULTS

### A. Data Processing

The data used in this research is crawled using the Twitter API, using an API that can be accessed via <https://developer.twitter.com/en>. The data from this crawl is a collection of opinions from Twitter users on the COVID-19 Vaccine for the period 01 May 2021.

The tweet data is crawled using several keywords such as "covid vaccine" and "corona vaccine" to find out how people think about the overall covid19 vaccine and "synovac vaccine", "astrazeneca vaccine", "Sinopharm vaccine" and

"Indonesian vaccine" to find out how public opinion on which type of vaccine is best. Each keyword is crawled at least 100 tweets which will be stored in the MySQL database.

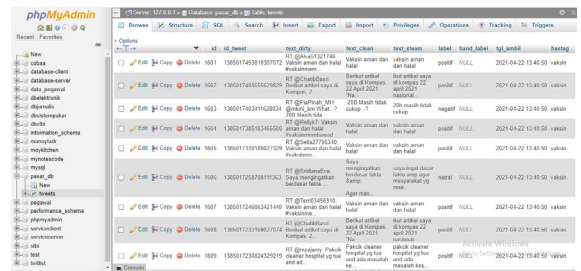


Figure 4. Database on MySQL

Each tweet data will be preprocessed for the purpose of removing noise and making it easier for later classification. There are two stages of preprocessing with Cleaning and Stemming.

```

//cleansing
$text_clean = new Clean($string);
$text_clean = $text_clean->toString();

//stem
$text_stem = $stemmer->stem($text_clean);
if(empty($text_stem)){
    $notIn++;
    break;
}
  
```

Figure 5. Source Code Cleaning and Stemming

The tweet data classification process is carried out using the naïve Bayes classifier algorithm by utilizing the lexicon dictionary which is sourced from the <https://github.com/masdevid/ID-OpinionWords> link. In the dictionary, there are more than 6,000 common words that have positive, negative and neutral sentiments.

The determination of tweet data is based on the overall accumulated value of words. If a tweet has a positive value it weights 0.333, if the value is negative it weights -0.333, and if the value is neutral it weights 0.334.

```

private $prior = array(
    'positif' => 0.3333333333333333,
    'negatif' => 0.3333333333333333,
    'netral' => 0.3333333333333334
);

public function __construct($dataFolder = false) {
    $this->setDataFolder($dataFolder);
    $this->loadDefaults();
}

public function score($sentence) {
    foreach ($this->negPrefixList as $negPrefix) {
        if (strpos($sentence, $negPrefix) !== false) {
            $sentence = str_replace($negPrefix, '.', $negPrefix, $sentence);
        }
    }

    $tokens = $this->getTokens($sentence);
    $total_score = 0;
    $scores = array();
  
```

Figure 6. Sentence weighting source code

Which are classified with the help of the naïve Bayes classifier algorithm.

```

public function categorise($sentence) {
    $scores = $this->score($sentence);
    $classification = key($scores);
    return $classification;
}
  
```

Figure 7. Source Code of Sentence Classification

### B. System Implementation

Using the PHP programming language and MySQL database storage, this web-based application system will connect directly to the Twitter API to retrieve tweets as public opinion, which will be used to measure positive, negative, and neutral sentiments.

#### 1. HOME



Figure 8. Home menu

In figure 6. The home menu contains a welcome welcome, information on the steps to carry out the analysis to the end and author information.

#### 2. DATA ANALYSIS

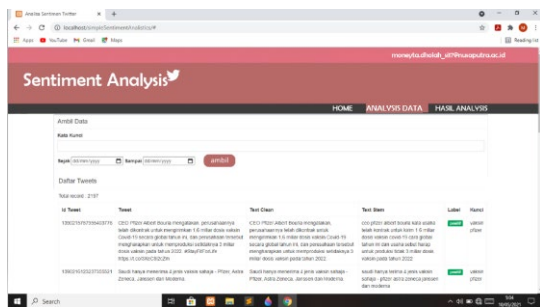


Figure 9. Data Analysis

Furthermore, in Figure 7, there is a Data Analysis display containing data crawling that is directly stored in the database and processed by text processing, cleaning and stemming, which then enters the labeling stage with the help of the naïve Bayes classifier method as a classification method.

#### 3. ANALYSIS RESULTS

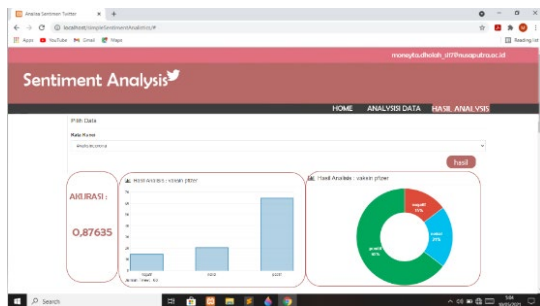


Figure 10. Analysis results

The last menu in this application is Figure 8. The results of the analysis contain the results of the previous analysis by visualizing them using bar charts and pie charts. Shows each of the previous data presentations and to be more convincing, the results of its accuracy are also displayed.

### C. Sentiment Analysis Results of the Covid-19 Vaccine

The results of the sentiment analysis surrounding the Covid-19 vaccine can be seen in table 3.

TABLE 3. RESULTS OF THE COVID-19 VACCINE SENTIMENT ANALYSIS

Keywords	Sentiment		
	Negative	Neutral	Positive
Covid vaccine	23%	31%	46%
Corona vaccine	49%	4%	47%
amount	72%	35%	93%

The Covid19 vaccine received 23% negative opinion, 31% neutral sentiment, and 46% good sentiment, while the Corona vaccination received 49% negative sentiment, 4% neutral sentiment, and 47% positive sentiment. According to the findings of the study, public opinion on this vaccination is overwhelmingly positive, with 93 percent of respondents agreeing, while negative and neutral responses are 72 percent and 35 percent, respectively.

### D. Results of Sentiment Analysis on the Type of Covid-19 Vaccine

The results of the sentiment analysis regarding the types of Covid-19 vaccines in Indonesia can be seen in table 4.

TABLE 4 SENTIMENT ANALYSIS RESULTS FOR COVID-19 VACCINE TYPES

Keywords	Sentiment		
	Negative	Neutral	Positive
Sinovac vaccine	57%	2%	41%
Astrazeneca vaccine	80%	6%	14%
Sinopharm vaccine	12%	4%	84%
Vaksin nusantaras	23%	6%	71%

From the table above, it can be concluded that the type of vaccine that has the most positive sentiment is Sinopharm vaccine at 84% and the smallest is the astrazeneca vaccine.

### E. Accuracy of Sentiment Analysis Results

Accuracy, is the ratio of true predictions (positive and negative) to the overall data [11]. In this study each of them has an accuracy as in table 5.

TABLE 5. ACCURACY RESULTS

Keywords	Sentiment			Accuracy
	Negative	Neutral	Positive	
Covid vaccine	23%	31%	46%	94,74%
Corona vaccine	49%	4%	47%	75,47%
Sinovac vaccine	57%	2%	41%	86,96%
Astrazeneca vaccine	80%	6%	14%	92,93%
Sinopharm Vaccine	13%	14%	73%	74,23%
Vaksin nusantaras	23%	6%	71%	71,43%

The "covid vaccine" is 94.74 percent accurate, the "corona vaccine" is 75.47 percent accurate, the "Sinovac vaccine" is 86.96 percent accurate, the "AstraZeneca vaccine" is 92.93 percent accurate, the "Sinopharm vaccine" is 74.23 percent accurate, and the "Nusantara vaccine" is 71.43 percent accurate. Each keyword has a high level of precision. The Covid Vaccine has a precision of 94.74 percent, the Corona Vaccine has a precision of 75.47 percent, the Sinovac Vaccine has a precision of 86.96 percent, the Astrazeneca Vaccine has a precision of 92.93 percent, the Sinopharm Vaccine has a precision of 74.23 percent, and the Nusantara Vaccine has a precision of 71.43 percent. This can be discovered using the Naive Bayes test formula, which is as follows:

$$P(P(NA)) = \frac{\text{Data}}{X} \times 100\%$$

Where:

$P(NA)$  : Percentage of accuracy value

Data : the number of test data that is true

$X$  : total sample size

As a result, the Naive Bayes algorithm does a good job of classifying.

## V. CONCLUSIONS

Based on the research that has been done, that the public sentiment, especially Twitter users, on the COVID-19 vaccine in 100 crawled tweet data, with positive results as much as 93%, negative 72% and neutral 35% with an accuracy of 94.74% for the keyword "covid vaccine". And 75.47% for the keyword "corona vaccine". Meanwhile, the types of vaccines circulating in Indonesia had the most positive results for the Sinopharm Vaccine as much as 84% with an accuracy of 74.23%, followed by the Nusantara Vaccine as much as 71% with an accuracy of 71.43%, the Sinovac Vaccine 41% with an accuracy of 86, 96% and the Astrazeneca Vaccine 14% with an accuracy of 92.93%.

## REFERENCES

- [1] "covid19com." <https://covid19.go.id/p/masyarakat-umum/3-vaksin-covid-19-dalam-program-vaksinasi-nasional-saat-ini-semua-direkomendasikan-who>.
- [2] A. Mustofa Hidayat and M. Syafrullah, "Algoritma Naïve Bayes Dalam Analisis Sentimen Untuk Klasifikasi Pada Layanan Internet PT.XYZ," *J. Telemat. MKOM*, vol. 9, no. 2, pp. 91–95, 2017, [Online]. Available: <http://journal.budiluhur.ac.id/index.php/telematika/article/view/532>.
- [3] H. Yulianton, F. A. Sutanto, and K. Hadiono, "Analisa Electronic Word of Mouth (E-Wom) Pada Media Sosial Twitter," *Dinamik*, vol. 22, no. 1, pp. 11–22, 2017, doi: 10.35315/dinamik.v22i1.7102.
- [4] L. Ardiani, H. Sujaini, and T. Tursina, "Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak," *J. Sist. dan Teknol. Inf.*, vol. 8, no. 2, pp. 44–51, 2020, doi: 10.26418/justin.v8i2.36776.
- [5] Azhar, S. U. Masruroh, L. K. Wardhani, and Okfalisa, "Perbandingan Kinerja Algoritma Naive Bayes Dan K-Nn Pendekatan Lexicon Pada Analisis Sentimen Di Media," *Pros. Semin. Nas. Fis. Univ. Riau IV*, no. September, pp. 978–979, 2019.
- [6] M. S. Hadna, P. I. Santosa, and W. W. Winarno, "Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2016, no. Sentika, pp. 57–64, 2016, [Online]. Available: <https://fti.uajy.ac.id/sentika/publikasi/makalah/2016/95.pdf>.
- [7] RI, Ministry of Health. 2020. "#INFOVAKSIN POCKET." In BUKU SAKU #INFOVAKSIN, by the Committee for Handling COVID-19 and Economic Recovery of the Republic of Indonesia. Jakarta: KOMINFO RI.
- [8] B. Liu, *Sentiment Analysis and Opinion Mining*. Morgan and Claypool Publishers, May 2012.
- [9] A. Ortigosa, J. M. Martin, and R. M. Carro, "Sentiment analysis in Facebook and its application to e-learning", *Computers in Human Behavior.*, vol. 31, pp. 527–541, 2014.
- [10] A. Erfina and Y. H. Putra, "Irony Sentence Detection Techniques Using Fuzzy Mamdani Historical Classifier", *IOP Conference Series: Materials Science and Engineering*, vol. 662, 2019.
- [11] L. K. O. S. Ni Putu Sri Merta Suryani, "Use of the Naïve Bayes Classifier Method in Analysis," *Using the Naïve Bayes Classifier Method in Analysis*, 2019.
- [12] L. D'Monte, "Swine Flu's Tweet Causes Online Flutter.," 2019. [http://www.business-standard.com/article/technology/swine-flu-s-tweet-tweet-causes-online-flutter-109042900097\\_1.html](http://www.business-standard.com/article/technology/swine-flu-s-tweet-tweet-causes-online-flutter-109042900097_1.html) (accessed Feb. 20, 2021).
- [13] N. Monarizqa, L. E. Nugroho, and B. S. Hantono, "Application of Sentiment Analysis on Indonesian Twitter as a Rating Officer," *J. Researcher. Tech. Electrical and Technol. Inf.*, Vol. 1, pp. 151–155, 2014.
- [14] Routray, Swain, and Mishra, "A Survey on Sentiment Analysis", *International Journal of Computer Applications*, August, Vol 70, Issue 10, 2013.
- [15] C. Troussas, M. Virvou, K. J. Espinosa, K. Llaguno, and J. Caro, "Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning", *IEEE.*, 2013.
- [16] Astaridewi, "News Reporting Classification System Using the Naive Bayes Method (Case Study of Government Official Sites). Jimbaran: Informatics Engineering Study Program, Computer Science Department, Faculty of Mathematics and Natural Sciences, Udayana University ". Thesis, Udayana University, Indonesia, 2015.
- [17] F. A. Irfan Wahyudin, Eneng Tita Tosida, *Theory and Practical Guide to Data Science and Big Data*, I. Indonesia: Institute for Research and Community Service, Pakuan University, 2019. indonesia: Lembaga Penelitian dan Pengabdian pada Masyarakat Universitas Pakuan, 2019.
- [18] H. C. Husada and A. S. Paramita, "Sentiment Analysis on Airlines on Twitter Platform Using Support Vector Machine (SVM) Algorithm", *Teknika*, vol. 10, no. 1, pp. 18-26, Feb. 2021.

