

# Analisis Sentimen Vaksinasi *Booster* Covid-19 Pada Platform Twitter Menggunakan Metode Naïve Bayes

Dessy Tri Anggraeni

Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Gunadarma, Bekasi, Indonesia

[dessytri@staff.gunadarma.ac.id](mailto:dessytri@staff.gunadarma.ac.id)

**ABSTRACT** – Since the end of 2019, the Covid-19 virus hit the whole world, including in Indonesia. One of the efforts to deal with the Covid-19 virus is vaccination. In Indonesia, the government requires people to vaccinate 3 times, that are First Vaccination, Second Vaccination, and Booster Vaccination. The public's response to the booster vaccine are varies. This study aims to reveal public sentiment towards booster vaccine activities. The research was conducted by collecting tweet data from the Twitter platform. The research was conducted by collecting data tweets from Twitter. The method used is the Naïve Bayes Classifier because the method is simple, the process is fast, and it has a fairly high level of confidence. In this method, public sentiment is divided into three, that are positive, neutral, and negative. The results showed that most people responded positively to this booster vaccine activity with a value of 56.8%, neutral as much as 39.9%, and negative as much as 3.3% with an accuracy rate of 86%.

**Keywords:** *Classification; Covid-19; Naïve Bayes Classifier; Sentiment; Twitter*

**ABSTRAK** – Sejak akhir tahun 2019, virus Covid-19 melanda seluruh dunia, tak terkecuali Indonesia. Salah satu upaya penanganan virus Covid-19 adalah dengan vaksinasi. Pemerintah Indonesia mewajibkan masyarakat untuk melakukan vaksinasi sebanyak tiga kali yaitu vaksinasi pertama, vaksinasi kedua, dan vaksinasi *booster*. Vaksinasi *booster* dilakukan ketika sebaran virus covid-19 sudah mulai terkendali sehingga respon masyarakat pun beragam. Penelitian ini ingin mengungkap sentimen masyarakat terhadap kegiatan pemberian vaksinasi *booster*. Penelitian dilakukan dengan mengumpulkan data *tweet* (cuitan) dari media sosial Twitter dan mengklasifikasikannya menggunakan metode Naïve Bayes Classifier. Metode ini dipilih karena metodenya yang sederhana, cepat, dan memiliki tingkat kepercayaan yang cukup tinggi. Pada metode ini sentimen masyarakat dibagi menjadi tiga yaitu positif, netral, dan negatif. Penelitian yang dihasilkan menunjukkan bahwa sebagian besar masyarakat merespon positif kegiatan vaksinasi *booster* dengan nilai sebesar 57%, netral sebanyak 40%, dan negatif sebanyak 3% dengan tingkat akurasi sebesar 86%.

**Kata Kunci:** *Covid-19; Klasifikasi; Naïve Bayes Classifier; Sentimen; Twitter.*

## 1. PENDAHULUAN

Virus Covid-19 mulai melanda dunia pada akhir tahun 2019, dimulai dari Kota Wuhan, Tiongkok [1]. WHO mengkategorikan Covid-19 sebagai pandemi dan hingga kini telah terjadi kurang lebih 632.533.408 kasus positif di seluruh dunia dengan 6.592.320 jumlah kematian [2]. Kasus Covid-19 di Indonesia tercatat sebanyak kurang lebih 6.556.627 kasus positif dengan 159.068 jumlah kematian [3]. Covid-19 merupakan penyakit yang disebabkan oleh virus *severe acute respiratory syndrome coronavirus 2* atau disingkat SARS-CoV-2. Virus ini menyerang bagian pernafasan, terutama paru – paru. Akibat yang dirasakan penderita antara lain sesak nafas, hilangnya kemampuan penciuman hingga sindrom pernafasan akut yang dapat memicu terjadinya kematian. Penularan virus yang begitu cepat sehingga membuat penanganannya menjadi sulit dilakukan. Selain melakukan pembatasan jarak social (social distancing), vaksinasi adalah cara yang dinilai ampuh memutus mata rantai penyebaran [1].

Pemerintah Indonesia mewajibkan seluruh masyarakat untuk melakukan vaksinasi sebanyak tiga kali yaitu vaksinasi ke-1, vaksinasi ke-2, dan vaksinasi booster. Saat ini capaian vaksinasi ke-1 adalah 87,47%, vaksinasi

ke-2 adalah 73,35%, dan vaksinasi booster adalah 28,03% [4]. Namun begitu, pro kontra pelaksanaan program vaksinasi terjadi di masyarakat, khususnya vaksinasi booster. Sebagian mendukung program ini dan sebagian yang lain menolak dengan berbagai alasan. Salah satu cara untuk mengetahui tanggapan masyarakat adalah melalui analisis sentimen. Analisis sentimen adalah teknik untuk mengekstrak informasi berupa pandangan terhadap suatu kejadian atau isu tertentu. Hasil analisis bisa berupa respon positif, negatif, maupun netral [5]. Hasil analisis bisa dijadikan dasar untuk evaluasi program dan dasar perbaikan di masa depan.

Pada penelitian ini digunakan metode Naive Bayes Classifier. Metode Naive Bayes merupakan salah satu metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik [6]. Algoritma ini dipilih karena metodenya yang sederhana, prosesnya cepat, dan tingkat kepercayaan yang cukup tinggi. Bahasa pemrograman yang digunakan adalah bahasa Python dengan memanfaatkan *library* yang sudah tersedia seperti *tweetblob*, *pandas*, dan *scikit-learn*.



### 2. DASAR TEORI

Penelitian yang berkaitan dengan sentimen masyarakat terhadap program vaksinasi sudah dilakukan oleh beberapa peneliti sebelumnya. Sri dan Sudi dalam penelitiannya yang berjudul “Analisis Sentimen Vaksin Sinovac Pada Twitter Menggunakan Algoritma Naïve Bayes” berhasil menerapkan algoritma Naïve Bayes untuk mengetahui sentimen masyarakat terhadap penggunaan vaksin Sinovac. Hasil penelitiannya adalah sentimen positif lebih besar dibandingkan sentimen negatif dengan tingkat akurasi 92,96% [6]. Penelitian serupa juga dilakukan Rima Tamara, dkk dengan hasil yang kurang lebih sama [7]. Zulfikar dan Nila juga melakukan penelitian serupa dengan menggunakan algoritma Naïve Bayes. Hasil penelitiannya adalah sentimen positif lebih besar daripada sentimen negatif untuk kegiatan vaksinasi [8].

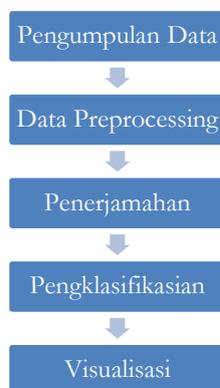
Hervinsyah dan Arita juga melakukan analisa sentimen pengguna Twitter terhadap kegiatan vaksinasi menggunakan algoritma SVM. Hasil penelitian menunjukkan bahwa kegiatan vaksinasi direspon positif dengan persentase sebanyak 83,6 % dengan tingkat akurasi 89% [9].

Sementara itu, Fitrah dan Kadek mencoba membandingkan kinerja metode Lexicon Based dan Naïve Bayes Classifier untuk mengukur sentimen masyarakat terhadap aplikasi investasi di Twitter. Hasilnya adalah Algoritma Naïve Bayes Classifier lebih akurat dibandingkan Lexicon Based dengan nilai 78% berbanding 67% [10].

### 3. METODOLOGI

Metode penelitian adalah gambaran langkah-langkah yang dilakukan dalam penelitian. Penggunaan metodologi akan membuat pelaksanaan penelitian dapat berjalan dengan lebih terencana dan terstruktur. Gambar 1 berikut merupakan metodologi dalam penelitian ini.

Tahapan pada penelitian ini terdiri dari pengumpulan data, *data preprocessing*, penerjemahan, pengklasifikasian, dan visualisasi.



Gambar 1. Tahapan Penelitian

Langkah pertama pada penelitian ini adalah pengumpulan data. Data diambil dari *public timeline tweet* yang dapat diakses dari situs netlytic.org. Netlytic adalah situs penganalisa teks dan jaringan sosial berbasis *cloud* yang secara otomatis dapat meringkas data tekstual dan menemukan jaringan komunikasi dari pos media sosial yang dapat diakses secara *public*. Pengambilan data dilakukan dengan menggunakan *filter* lokasi dan kata kunci. Lokasi yang dipilih adalah Indonesia dan kata kunci yang digunakan adalah vaksinasi *booster*, *booster*, dan vaksinasi ketiga. Data diambil dari tanggal 17 April 2022 sampai dengan 31 April 2022.

Selanjutnya data yang sudah dikumpulkan akan diseleksi agar bisa digunakan untuk pemrosesan lebih lanjut. Proses ini disebut dengan *preprocessing* [11]. Terdapat beberapa tahapan dalam proses ini, antara lain adalah data cleaning, case folding, stopword removal, tokenizing, dan stemming [11]. Secara singkat, data cleaning bisa dijelaskan dengan membersihkan data dari karakter/symbol/kata yang tidak akan digunakan. Sementara case folding adalah proses mengubah huruf capital menjadi huruf kecil atau sebaliknya dengan tujuan agar kalimat menjadi standar. Stopword Removal atau filtering adalah proses menghilangkan kata yang tidak sesuai atau tidak memiliki makna di dalam kalimat. Selanjutnya tokenizing adalah proses memecah sebuah kalimat menjadi kata per kata sebagai persiapan untuk proses selanjutnya yaitu stemming. Stemming adalah proses menentukan bentuk dasar dari sebuah kata.

Setelah proses *preprocessing* selesai dilakukan., langkah berikutnya adalah penerjemahan data. Data yang sudah melalui *preprocessing* dianggap menjadi data yang sudah bersih dan siap digunakan. Akan tetapi dikarenakan *library* yang digunakan hanya mendukung data berbahasa inggris, maka data *tweet* yang berbahasa Indonesia perlu diterjemahkan terlebih dahulu.

Setelah data siap digunakan, maka langkah selanjutnya adalah pengklasifikasian. Proses ini adalah proses utama dalam penelitian ini. Pada langkah ini data akan dikelompokkan atau diklasifikasikan berdasarkan label sentimen yang ditentukan yaitu positif, negatif, dan netral. Algoritma yang digunakan untuk klasifikasi adalah *Naïve Bayes Classifier*. Pada algoritma *Naïve Bayes*, penetapan label didasarkan pada perhitungan probabilitas dari rumus *Naïve Bayes*. Pada penelitian ini, perhitungan dilakukan secara otomatis oleh Python dengan menggunakan library TextBlob dan Scikit-Learn.

#### Algoritma Naïve Bayes

*Naïve Bayes* adalah teknik untuk klasifikasi yang populer karena mudah digunakan. Selain itu algoritma *Naïve Bayes* juga memiliki tingkat efektivitas yang tinggi.

Algoritma ini banyak digunakan pada bidang *data mining* dan *machine learning* [5]. Kelebihan lain dari penggunaan metode algoritma *Naïve Bayes* yaitu hanya memerlukan jumlah *data training* yang kecil atau sedikit.

Persamaan Teorema Bayes adalah sebagai berikut :

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Pada persamaan ini, A dan B merupakan kejadian.  $P(A|B)$  berarti peluang kejadian A jika kondisi B terjadi.  $P(B|A)$  berarti peluang kejadian B jika kondisi A terjadi.  $P(A)$  dan  $P(B)$  masing-masing adalah peluang kejadian A dan peluang kejadian B.

Langkah terakhir pada penelitian adalah visualisasi. Visualisasi dilakukan untuk memberikan gambaran secara visual mengenai persentase dari tiap sentimen yang didapat dari hasil klasifikasi. Visualisasi sentimen akan ditampilkan dalam bentuk *pie diagram*. Proses ini menggunakan library *matplotlib*.

#### 4. HASIL DAN PEMBAHASAN

Langkah-langkah pada penelitian ini dilakukan sesuai dengan alur yang telah dibuat sebelumnya pada bagian metodologi penelitian. Berikut adalah pembahasan lebih detail di tiap-tiap tahapan.

##### Pengumpulan Data

Pengumpulan data dilakukan dengan mengambil data dari situs *netlytic.org* dengan menggunakan *filter* lokasi dan kata kunci. Lokasi yang dipilih adalah Indonesia dan kata kunci yang digunakan adalah *vaksin booster*, *booster*, dan *vaksin ketiga*. Data diambil dari tanggal 17 April 2022 sampai dengan 31 April 2022. Jumlah data yang diperoleh sebanyak 2.500 *tweet* dalam format *.csv*. Gambar 2 adalah contoh data yang diambil dalam bentuk *.csv*.

Selanjutnya data *.csv* tersebut diubah menjadi bentuk *dataframe* agar lebih mudah untuk diproses. Pembuatan *dataframe* bisa dilakukan dengan menggunakan *library* *pandas*. Atribut yang diambil hanya dua, yaitu *author* yang merupakan akun pemilik *tweet* dan *tweet* yang merupakan isi *tweet*. Sementara atribut yang lain tidak diambil karena tidak diperlukan dalam proses klasifikasi. *Dataframe* yang terbentuk bisa dilihat pada *screenshot* di gambar 3 berikut.

	author	tweet
0	yashadipura	Anak 18 tahun ke bawah, dibebaskan dari syarat ...
1	KramatJa3	GIAT AKSELERASI VAKSINASI BOOSTER RAMADHAN DI ...
2	kaka_jadu	Bagi Pemudik yang sudah vaksinasj dosis ketiga...
3	mpratne	@dodon_jerry Mpo sudah divaksin booster tapi...
4	Ardinova7	Yang mudik jangan lupa vaksin booster, nanti di...
...	...	...
2495	daandailamy	@daandailamy Selamat sore Kak. Bila sudah mela...
2496	LantasDairi	Personil "Unit Turjawali" Sat Lantas berserta...
2497	AnonFans62	Seluruh umat Islam Indonesia harus sadar bahwa...
2498	RatuZulfaz	Viral Polisi Larang Warga Terima Bantuan karena...
2499	HUMASPolrespyk	Mudik..? Ayo Lengkapi Vaksinasi Boosternya...

Gambar 2. Data Setelah Diubah Menjadi *Dataframe*.

##### Preprocessing

Pada tahap ini dilakukan beberapa tahapan pembersihan data yaitu *data cleaning*, *case folding*, *stopwords removal*, *tokenization*, dan *stemming*.

##### 1) Data Cleaning

Pada tahap *data cleaning* dilakukan proses pembersihan data dari karakter yang tidak digunakan. Contohnya adalah menghilangkan nama pengguna, data yang terduplikasi, data kosong, tanda koma, tanda seru, tanda titik, tanda tanya, link URL, "@", menghapus emoji, *hashtag*, dan *mention*. Hasil dari *data cleaning* bisa dilihat pada gambar 4 berikut.

	tweet	cleaning
0	Anak 18 tahun ke bawah, dibebaskan dari syarat ...	Anak tahun ke bawah dibebaskan dari syarat ...
1	GIAT AKSELERASI VAKSINASI BOOSTER RAMADHAN DI ...	GIAT AKSELERASI VAKSINASI BOOSTER RAMADHAN DI ...
2	Bagi Pemudik yang sudah vaksinasj dosis ketiga...	Bagi Pemudik yang sudah vaksinasj dosis ketiga...
3	@dodon_jerry Mpo sudah divaksin booster tapi...	Mpo sudah divaksin booster tapi...
4	Yang mudik jangan lupa vaksin booster, nanti di...	Yang mudik jangan lupa vaksin booster nanti di...
...	...	...
2495	@daandailamy Selamat sore Kak. Bila sudah mela...	Selamat sore Kak Bila sudah mela...
2496	Personil "Unit Turjawali" Sat Lantas berserta...	Personil Unit Turjawali Sat Lantas berserta...
2497	Seluruh umat Islam Indonesia harus sadar bahwa...	Seluruh umat Islam Indonesia harus sadar bahwa...
2498	Viral Polisi Larang Warga Terima Bantuan karena...	Viral Polisi Larang Warga Terima Bantuan karena...
2499	Mudik..? Ayo Lengkapi Vaksinasi Boosternya...	Mudik Ayo Lengkapi Vaksinasi Boosternya...

Gambar 3. Data *Tweet* Setelah Melalui Proses *Data Cleaning*.

##### 2) Case Folding

*Case folding* merupakan tahapan untuk mengubah huruf kapital menjadi huruf kecil atau sebaliknya. Hasil dari tahapan ini bertujuan agar dokumen yang akan diolah memiliki bentuk standar. Contoh proses *case folding* bisa dilihat pada gambar 5 berikut.

cleaning	case folding
Anak tahun ke bawah dibebaskan dari syarat ...	anak tahun bawah dibebaskan syarat ...
GIAT AKSELERASI VAKSINASI BOOSTER RAMADHAN DI ...	giat akselerasi vaksinasi booster ramadhan ...
Bagi Pemudik yang sudah vaksinasi dosis ketiga...	pemudik sudah vaksinasi dosis ketiga...
Mpo sudah divaksin booster tapi...	mpo divaksin booster ...
Yang mudik jangan lupa vaksin booster nanti di...	mudik lupa vaksin booster ...
...	...
Selamat sore Kak Bila sudah mela...	selamat sore kak bila mela...
Personil Unit Turjawali Sat Lantas berserta...	personil unit turjawali sat lantas berserta...
Seluruh umat Islam Indonesia harus sadar bahwa...	seluruh umat islam indonesia sadar bahwa...
Viral Polisi Larang Warga Terima Bantuan karena...	viral polisi larang warga terima bantuan...
Mudik Ayo Lengkapi Vaksinasi Boosternya...	mudik lengkapi vaksinasi boosternya...

Gambar 4. Data *Tweet* Setelah Melalui *Case Folding* dan *Stopword Removal*.

### 3) Stopword Removal

*Stopword removal* adalah proses menghilangkan kata yang tidak memiliki makna atau kata yang tidak sesuai dengan

id	tweetid	author	description	pubdate	source	favorite_counter	retweet_counter	language
0	15181504866	yashadipura	Anak 18 tahun ke bawah, dibebaskan dari syarat vaksinasi booster	24/04/2022	Twitter for Android	1	41	Indonesia
1	15181504473	KramatJa3	GIAT AKSELERASI VAKSINASI BOOSTER RAMADHAN DI JABODETABEK	24/04/2022	Twitter for Android	0	0	Indonesia
2	15181504799	kaka_jadu	Bagi Pemudik yang sudah vaksinasi dosis ketiga boleh melakukan perjalanan mudik	24/04/2022	Twitter for Android	0	1	Indonesia
3	15181504621	mporatne	@dodon_jerry Mpo sudah divaksin booster tapi masih kena covid juga	24/04/2022	Twitter for Android	0	1	Indonesia
4	15181496061	Ardinova7	Yang mudik jangan lupa vaksin booster, nanti disuruh balik lagi	24/04/2022	Twitter for Android	0	0	Indonesia
5	15181495103	daandailamy	@daandailamy Selamat sore Kak. Bila sudah melakukan vaksin booster, apakah sudah boleh mudik?	24/04/2022	Twitter for Android	0	1	Indonesia
6	15181483725	LantasDairi	Personil "Unit Turjawali" Sat Lantas berserta jajaran mengucapkan Selamat Melakukan Perjalanan Mudik	24/04/2022	Twitter for Android	0	0	Indonesia
7	15181481518	AnonFans62	Seluruh umat Islam Indonesia harus sadar bahwa ini adalah perjuangan bersama melawan Covid19	24/04/2022	Twitter for Android	0	0	Indonesia
8	15181478064	RatuZulfaz	Viral Polisi Larang Warga Terima Bantuan karena tidak masuk daftar keluarga tidak mampu	24/04/2022	Twitter for Android	0	0	Indonesia
9	15181458369	HUMASPolrespyk	Mudik...? Ayo Lengkapi Vaksinasi Boosternya...	24/04/2022	Twitter for Android	0	0	Indonesia

Gambar 6. Data yang di kumpulkan melalui *website* netlytic.org dalam Bentuk .csv.

### 4) Tokenizing

Proses *tokenizing* akan memecah sebuah kalimat atau *tweet* ke dalam bentuk kata per kata.

### 5) Stemming

Tahapan *stemming* merupakan proses mengubah kata menjadi bentuk dasarnya. Pada proses ini juga dilakukan proses menghapus imbuhan seperti awalan, sisipan, atau akhiran. Gambar 6. merupakan contoh *tweet* yang telah melewati tahap *tokenizing* dan *stemming*.

topik dokumen. Pada tahap ini juga dihilangkan beberapa kata dan kata singkatan yang sering ditemukan pada data *twitter* seperti "btw", "y", "wkwk", "dr" dan lainnya agar data menjadi lebih bersih. Hasil dari proses *case folding* dan *stopword removal* bisa dilihat pada gambar 6 berikut.

case folding	stop removal
anak tahun bawah dibebaskan syarat ...	anak tahun bawah dibebaskan syarat ...
giat akselerasi vaksinasi booster ramadhan ...	giat akselerasi vaksinasi booster ramadhan ...
pemudik sudah vaksinasi dosis ketiga...	pemudik sudah vaksinasi dosis ketiga...
mpo divaksin booster ...	mpo divaksin booster ...
mudik lupa vaksin booster ...	mudik lupa vaksin booster ...
...	...
selamat sore kak bila mela...	selamat sore kak bila mela...
personil unit turjawali sat lantas berserta...	personil unit turjawali sat lantas berserta...
seluruh umat islam indonesia sadar bahwa...	seluruh umat islam indonesia sadar bahwa...
viral polisi larang warga terima bantuan...	viral polisi larang warga terima bantuan...
mudik lengkapi vaksinasi boosternya...	mudik lengkapi vaksinasi boosternya...

Gambar 5. Data *Tweet* Setelah Melalui *Case Folding* dan *Stopword Removal*.

Setelah melalui tahapan *preprocessing* ini, tidak semua data *tweet* bisa digunakan. Dari 2500 data yang dikumpulkan, data yang bisa digunakan hanya berjumlah 1.678 data.

tokenizing	stemming
[anak, tahun, bawah, dibebaskan, syarat ...	anak tahun bawah bebas syarat ...
[giat, akselerasi, vaksinasi, booster, ramadhan ...	giat akselerasi vaksin booster ramadhan ...
[pemudik, sudah, vaksinasi, dosis, ketiga...	mudik sudah vaksinasi dosis tiga...
[mpo, divaksin, booster ...	mpo divaksin booster ...
[mudik, lupa, vaksin, booster ...	mudik lupa vaksin booster ...
...	...
[selamat, sore, kak, bila, mela...	selamat sore kak bila mela...
[personil, unit, turjawali, sat, lantas, berserta...	personil unit turjawali sat lantas serta...
[seluruh, umat, islam, indonesia, sadar, bahwa...	seluruh umat islam indonesia sadar bahwa...
[viral, polisi, larang, warga, terima, bantuan...	viral polisi larang warga terima bantuan...
[mudik, lengkapi, vaksinasi, boosternya...	mudik lengkap vaksin booster...

**Gambar 7.** Data setelah melewati tahap *tokenizing* dan *stemming*.

### Penerjemahan Data

Setelah data dibersihkan pada tahap *preprocessing*, selanjutnya data akan melalui proses penerjemahan data. Penerjemahan data dilakukan karena *library* yang digunakan hanya mendukung data berbahasa inggris. Proses penerjemahan dilakukan menggunakan *library google\_trans\_new*. Gambar 8. merupakan hasil sebelum dan sesudah melalui tahap penerjemahan data.

stemming	translate
anak tahun bawah bebas syarat ...	Underage children are free ...
giat akselerasi vaksin booster ramadhan ...	actively accelerating the Ramadan booster vaccine ...
mudik sudah vaksinasi dosis tiga...	I've been vaccinated on dose three...
mpo divaksin booster ...	mpo divaksin booster ...
mudik lupa vaksin booster ...	Homecoming forgot the booster vaccine ...
...	...
selamat sore kak bila mela...	good afternoon sis if doing...
personil unit turjawali sat lantas serta...	turjawali unit personnel sat directly as well...
seluruh umat islam indonesia sadar bahwa...	all Indonesian Muslims are aware of that...
viral polisi larang warga terima bantuan...	viral police prohibit residents from receiving assistance...
mudik lengkap vaksin booster...	homecoming complete with booster vaccines...

**Gambar 8.** Data yang Telah Diterjemahkan ke Bahasa Inggris.

### Pengklasifikasian

Pengklasifikasian data dilakukan untuk untuk menetapkan sebuah kalimat/*tweet* apakah bersentimen positif, netral, atau negatif. Proses pengklasifikasian menggunakan metode *Naive Bayes Classifier* dengan bantuan *library* TextBlob dan sk-learn. Hasil klasifikasi dapat dilihat pada gambar 9.

translate	sentiment
Underage children are free ...	Positif
actively accelerating the Ramadan booster vaccine ...	Positif
I've been vaccinated on dose three...	Netral
mpo divaksin booster ...	Netral
Homecoming forgot the booster vaccine ...	Netral
...	...
good afternoon sis if doing...	Netral
turjawali unit personnel sat directly as well...	Positif
all Indonesian Muslims are aware of that...	Positif
viral police prohibit residents from receiving assistance...	Netral
homecoming complete with booster vaccines...	Positif

**Gambar 9.** Data yang Sudah Diklasifikasi Menggunakan *Naive Bayes Classifier*

Setelah data diklasifikasi, hasilnya terdapat 953 data bersentimen positif, 669 data netral, dan 56 data bersentimen negatif, dengan total data sebanyak 1678 data. Sehingga, bisa disimpulkan bahwa respon masyarakat terhadap vaksinasi *booster* adalah positif.

```

Hasil Analisis Data:
Positif = 953
Netral = 669
Negatif = 56
Total Data : 1678
    
```

**Gambar 10.** Hasil Pengklasifikasian Data.

Penghitungan akurasi juga dilakukan menggunakan *Naive Bayes Classifier*. Hasil perhitungan menunjukkan bahwa tingkat akurasi klasifikasi adalah sebesar 0,86 atau 86%. Gambar 11 merupakan hasil perhitungan akurasi menggunakan *Naive Bayes Classifier*.

```

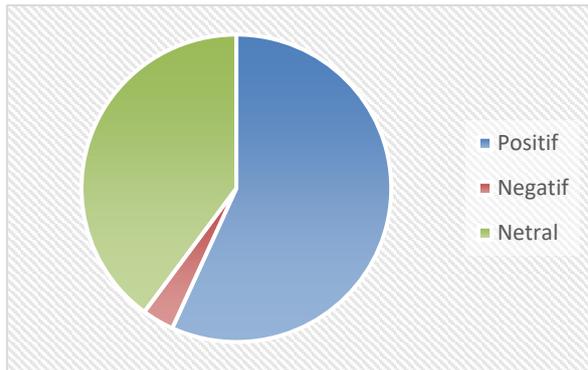
from textblob.classifiers import NaiveBayesClassifier
cl = NaiveBayesClassifier(train_set)
print('Akurasi Test:', cl.accuracy(dataset))

Akurasi Test: 0.8617401668653158
    
```

**Gambar 11.** Pengukuran akurasi.

### Visualisasi

Data sentimen yang telah didapatkan dari proses sebelumnya, kemudian ditampilkan dalam bentuk visual menggunakan *pie diagram*. Diagram pada gambar 12 berikut merupakan penggambaran dari hasil pengklasifikasian data *tweet*.



**Gambar 12.** Tampilan *Pie Diagram* Sentimen Masyarakat Terhadap Vaksinasi *Booster*.

Visualisasi yang dihasilkan menunjukkan jumlah sentimen positif 57%, sentimen negatif 3%, dan sentimen netral 40%. Hasil visualisasi ini membuktikan bahwa secara umum masyarakat menerima secara positif program vaksinasi *booster* Covid-19.

### 5. KESIMPULAN

Penelitian untuk mengetahui sentimen masyarakat terhadap kegiatan vaksinasi *booster* ini berhasil dilakukan. Penelitian melibatkan sebanyak 2.500 data *tweet* dimana setelah dilakukan proses *preprocessing* hanya 1678 data yang bisa diproses ke tahapan klasifikasi. Pengklasifikasian dilakukan menggunakan algoritma *Naïve Bayes Classifier* dan didapatkan data bahwa sebagian besar masyarakat merespon secara positif kegiatan vaksinasi *booster*. Presentase sentimen menunjukkan bahwa sebanyak 56,8 % merespon positif kegiatan vaksinasi *booster*, sebanyak 39,9 % memilih netral, dan sisanya sebanyak 3,3 % merespon negatif.

### DAFTAR PUSTAKA

- [1] A. Susilo and C. M. Rumende, "Coronavirus Disease 2019: Tinjauan Literatur Terkini," *Jurnal Penyakit Dalam Indonesia*, vol. 7, no. 1, 2020.
- [2] W. H. Organization, "WHO Coronavirus Dashboard," World Health Organization, [Online]. Available: covid19.who.int. [Accessed 8 November 2022].
- [3] C.-1. Hotline, "Situasi COVID-19 di Indonesia (Update per 12 November 2022)," covid19.go.id, [Online]. Available: https://covid19.go.id/artikel/2022/11/12/situasi-covid-19-di-indonesia-update-12-november-2022. [Accessed 12 November 2022].
- [4] C.-1. Hotline, "Situasi Covid-19 Nasional," covid19.go.id, [Online]. Available: https://covid19.go.id/situasi. [Accessed 12 November 2022].
- [5] P. A. Sumitro, D. I. Mulyana and W. Saputro, "Analisis Sentimen Terhadap Vaksin Covid-19 di Indonesia pada Twitter Menggunakan Metode Lexicon Based," *Jurnal Informatika dan Teknologi Komputer*, vol. 2, no. 2, pp. 50-56, 2021.
- [6] S. Lestari and S. Saepudin, "Analisis Sentimen vaksin Sinovac Pada Twitter Menggunakan Algoritma Naïve Bayes," *SISMATIK*, vol. 1, no. 1, 2021.
- [7] A. R. T. Aldisa and M. Aldinugroho, "Analisis Sentimen Megenai Vaksin Sinovac di Media Sosial Twitter Menggunakan Naïve Bayes Classification," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 6, no. 3, 2022.
- [8] Z. Firmansyah and N. F. Puspitasari, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Berdasarkan Opini Pada Twitter Menggunakan Algoritma Naïve Bayes," *Jurnal Teknik Informatika*, vol. 14, no. 2, 2021.
- [9] Herwinsyah and A. Witanti, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Jurnal Sistem Informasi dan Informatika*, vol. 5, no. 1, 2022.
- [10] F. Amaliah and K. D. Nuryana, "Perbandingan Akurasi Metode Lexicon Based dan Naïve Bayes Classifier Pada Analisis Sistem Sentimen Pendapat Masyarakat Terhadap Aplikasi Investasi Pada Media Twitter," *Journal of Informatics and Computer Science*, vol. 3, no. 3, 2022.
- [11] S. Ambiyar, R. Watrianthos and F. E. Verawardina, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes," *J. Media Inform. Budidarma*, vol. 5, pp. 157-163, 2021.