

**SKRIPSI**

**ANALISIS SENTIMEN PENGGUNA TWITTER  
TERHADAP VAKSIN COVID-19 DI INDONESIA  
MENGUNAKAN METODE *NAÏVE BAYES*  
*CLASSIFIER***



*Unipa Surabaya*

**Yuni Mega Lestari  
NIM. 172400005**

**PROGRAM STUDI STATISTIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS PGRI ADI BUANA SURABAYA  
2022**

## LEMBAR PERSETUJUAN

Yang bertanda tangan dibawah ini menyatakan bahwa:

Nama : Yuni Mega Lestari  
NIM : 172400005  
Pembimbing 1 : Artanti Indrasetianingsih, S.Si., M.Si.  
Pembimbing 2 : Muhammad Athoillah, S.Si., M.Si.  
Judul Skripsi : Analisis Sentimen Pengguna Twitter  
terhadap Vaksin Covid-19 Di Indonesia  
Menggunakan Metode *Naïve Bayes  
Classifier*

Disetujui untuk diseminarkan pada seminar skripsi di Program Studi  
Statistika Universitas PGRI Adi Buana Surabaya.

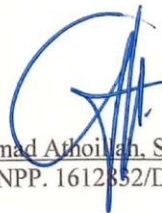
Surabaya, 20 Juni 2022

Pembimbing 1



Artanti Indrasetianingsih, S.Si., M.Si.  
NPP. 0609466/DY

Pembimbing 2




Muhammad Athoillah, S.Si., M.Si.  
NPP. 1612852/DY


## LEMBAR PENGESAHAN

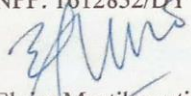
Skripsi disusun untuk memenuhi salah satu syarat memperoleh gelar  
Sarjana Statistika (S.Stat.)  
Di Universitas PGRI Adi Buana Surabaya  
Oleh :

Nama : Yuni Mega Lestari  
NIM : 172400005  
Pembimbing 1 : Artanti Indrasetyaningih, S.Si., M.Si.  
Pembimbing 2 : Muhammad Athoillah, S.Si., M.Si.  
Judul Skripsi : Analisis Sentimen Pengguna Twitter terhadap  
Vaksin Covid-19 Di Indonesia dengan  
*Naïve Bayes Classifier*

Tanggal Ujian : Kamis, 30 Juni 2022  
Disetujui Oleh Tim Penguji Skripsi :

  
Artanti Indrasetyaningih, S.Si., M.Si. (Pembimbing 1)  
NPP. 0609466/DY

  
Muhammad Athoillah, S.Si., M.Si. (Pembimbing 2)  
NPP. 1612832/DY

  
Elvira Mustikawati P.H. S.Si., M.Si. (Penguji)  
NPP. 1610820/DY

Mengesahkan  
  
Alfisyahrina Hapsery, S.Si., M.Si.  
NPP. 1804856/DY



## BERITA ACARA UJIAN SKRIPSI

Skripsi ini diterima dan disetujui oleh panitia ujian skripsi Sarjana Statistika Program Studi Statistika Fakultas Sains dan Teknologi Universitas PGRI Adi Buana Surabaya.

Pada Hari : Kamis  
Tanggal : 30 Juni  
Tahun : 2022

Panitia Ujian Skripsi :

1. Ketua Dekan :



  
Dra. Diah Karunia Binawati, M.Si.

2. Sekretaris :

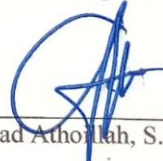


Dr. Arif Yachya, S.Si., M.Si.

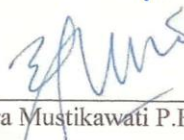
3. Anggota :



Artanti Indrasetianingsih, S.Si., M.Si.



Muhammad Athollah, S.Si., M.Si.



Elvira Mustikawati P.H, S.Si., M.Si.

## KATA PENGANTAR

Puji syukur penulis ucapkan kehadiran Allah SWT yang telah melimpahkan rahmat serta hidayah-Nya sehingga penulis mampu menyelesaikan penelitian yang berjudul **“ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP VAKSIN COVID-19 DI INDONESIA DENGAN METODE NAÏVE BAYES CLASSIFIER”**. Penulis mengucapkan terima kasih kepada beberapa pihak yang secara langsung maupun tidak langsung telah membantu penyusunan skripsi.

1. Bapak Dr. Marianus Subandowo MS selaku Rektor Universitas PGRI Adi Buana Surabaya.
2. Ibu Dra. Diah Karunia Binawati, M.Si., selaku Dekan Fakultas Sains dan Teknologi yang telah memberikan bimbingan, arahan, motivasi dan semangat selama saya menjadi mahasiswa.
3. Ibu Artanti Indrasetianingsih, S.Si., M.Si., selaku dosen pembimbing dan Wakil Dekan II Fakultas Sains dan Teknologi yang telah membimbing, memberikan saran-saran, support, dan mengarahkan dengan sabar untuk dapat menyelesaikan skripsi ini.
4. Ibu Alfisyahrina Hapsery, S.Si., M.Si., selaku Ketua Program Studi Statistika yang telah memberikan pengarahan, motivasi dan pengajaran tentang ilmu statistika.
5. Ibu Elvira Mustikawati P.H, S.Si., M.Si., selaku dosen penguji yang telah memberikan saran-saran yang membangun untuk kesempurnaan skripsi ini.
6. Bapak Muhammad Athoillah, S.Si., M.Si., selaku dosen pembimbing yang telah memberikan bimbingan, saran dan arahan untuk dapat menyelesaikan skripsi ini.
7. Seluruh dosen dan tenaga pendidik Program Studi Statistika yang telah memberikan ilmu selama penulis menempuh pendidikan, serta membantu kelancaran dan kemudahan dalam pelaksanaan kegiatan perkuliahan.
8. Teman-teman Mahasiswa Statistika yang tidak dapat disebutkan satu per satu oleh penulis yang telah memberi bantuan untuk penulis dalam perkuliahan dan dalam menyelesaikan skripsi ini.

9. Orang penulis (Bapak Sugiyanto dan Ibu Nanik Suliyani), saudara-saudari penulis yang telah memberikan kasih sayang, dukungan, motivasi, nasihat dan doa.
10. Alviandra Satya Dwi Putra yang telah memberikan dukungan, saran, motivasi, nasihat dan doa.
11. Teman-teman Pondok Pesantren Mahasiswa Khoirul Huda 3 & 1 yang tidak dapat disebutkan satu per satu oleh penulis yang telah memberi bantuan, saran, support dan doa.
12. Semua pihak yang membantu dalam proses penyusunan skripsi ini.

Penulis menyadari bahwa laporan skripsi ini masih jauh dari kata sempurna, oleh karena itu penulis sangat mengharapkan kritik dan saran yang membangun agar dapat menjadi bahan untuk perbaikan berikutnya.

Surabaya, 9 Juni 2022

Penulis

## MOTTO

“Mencari Jalannya Syukur”

karena,

" لَعْنَةُ شُكْرِكُمْ لَا زَيْدَنَّكُمْ وَلَعْنُ كَفْرِكُمْ إِنَّ عَذَابِي لَشَدِيدٌ " سورة ابراهيم : ٧

“Jika kamu bersyukur, niscaya aku (Allah) akan menambah (nikmat) kepadamu, dan jika kamu mengingkari (nikmat-Ku), sesungguhnya siksa-Ku sangat berat” (Q.S. Ibrahim:7)

## PERSEMBAHAN

Saya persembahkan skripsi saya ini untuk kedua orangtua tercinta, papa dan ibu yang senantiasa memberikan cinta kasih, doa, penyemangat, motivasi, nasihat kepada saya, sehingga saya dapat menyelesaikan skripsi ini dengan baik. Terima kasih juga kepada saudara-saudaraku yang memberikan support, doa dan nasihat, juga kepada teman-teman PPM (Pondok Pesantren Mahasiswa) Khoirul Huda 3 dan 1 atas support dan doanya untuk saya. Tak lupa juga saya ucapkan terima kasih kepada ibu bapak dosen statistika yang telah mengajarkan ilmu statistika kepada saya, serta teman-teman dari prodi statistika semoga dapat bertemu lagi pada kesempatan lain.

## SURAT PERNYATAAN

Yang bertanda tangan di bawah ini, menyatakan bahwa :

Nama : Yuni Mega Lestari  
NIM : 172400005  
Judul Skripsi : Analisis Sentimen Pengguna Twitter  
terhadap Vaksin Covid-19 Di Indonesia  
Menggunakan Metode *Naïve Bayes*  
*Classifier*

Menyatakan bahwa skripsi ini adalah hasil karya saya dan bukan merupakan duplikasi sebagian atau seluruhnya dari karya orang lain, kecuali bagian sumber informasi yang dicantumkan.

Pernyataan ini dibuat dengan sebenar-benarnya secara sadar dan bertanggung jawab dan saya bersedia menerima sanksi pembatalan skripsi apabila terbukti melakukan duplikasi terhadap skripsi atau karya ilmiah lain yang sudah ada.

Surabaya, 9 September 2022



Yuni Mega Lestari



## DAFTAR ISI

HALAMAN JUDUL .....	i
LEMBAR PERSETUJUAN .....	ii
LEMBAR PENGESAHAN .....	iii
BERITA ACARA UJIAN SKRIPSI .....	iv
KATA PENGANTAR.....	v
HALAMAN PERSEMBAHAN.....	vii
SURAT PERNYATAAN .....	viii
ABSTRAK.....	ix
<i>ABSTRACT</i> .....	x
DAFTAR ISI .....	xi
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR .....	xiv
BAB I PENDAHULUAN .....	1
1.1    Latar Belakang .....	1
1.2    Rumusan Masalah .....	6
1.3    Tujuan Penelitian.....	6
1.4    Manfaat Penelitian.....	6
1.5    Batasan Masalah.....	7
BAB 2 TINJAUAN PUSTAKA .....	8
2.1 <i>Sentiment Analysis</i> .....	8
2.2 <i>Machine Learning</i> .....	9
2.3 <i>Text Mininng</i> .....	10
2.4    Praproses Teks.....	10
2.5 <i>Term Frequency</i> .....	16
2.6 <i>K-fold Cross Validation</i> .....	16

2.7	<i>Naïve Bayes Classifier</i> .....	17
2.8	Evaluasi Performa Klasifikasi .....	22
2.9	<i>Wordcloud</i> .....	23
2.10	COVID-19.....	24
2.11	Vaksin Covid-19.....	24
2.12	Twitter .....	25
BAB 3 METODOLOGI PENELITIAN .....		26
3.1	Sumber Data.....	26
3.2	Struktur Data .....	26
3.3	Langkah-langkah Penelitian.....	28
BAB 4 ANALISIS DAN PEMBAHASAN .....		32
4.1	Pengambilan Data.....	32
4.2	Praproses Teks.....	33
4.3	Karakteristik Data.....	34
4.4	Klasifikasi Menggunakan <i>Naïve Bayes Classifier</i> .....	40
4.6	Visualisasi <i>Word Cloud</i> .....	46
BAB 5 KESIMPULAN DAN SARAN .....		48
5.1	Kesimpulan .....	48
5.2	Saran.....	48
DAFTAR PUSTAKA.....		49
LAMPIRAN .....		52

## DAFTAR TABEL

Tabel 2.1 Hasil <i>Case Folding</i> Praproses Teks .....	12
Tabel 2.2 Hasil <i>Tokenizing</i> Praproses Teks .....	13
Tabel 2.3 Hasil <i>Stemming</i> Praproses Teks .....	14
Tabel 2.4 Hasil <i>Stopword</i> Praproses Teks .....	15
Tabel 2.5 <i>Confusion Matrix</i> .....	22
Table 3.1 Struktur Data Sebelum Praproses Teks.....	26
Tabel 3.2 Contoh Data Tweet Vaksin Covid-19 .....	27
Tabel 3.3 Struktur Data Setelah Praproses Teks .....	28
Tabel 4.1 Data Praproses Teks.....	33
Tabel 4.2 Data Tweet Vaksin yang telah dilabeli.....	35
Tabel 4.3 Data yang telah berbentuk <i>Document Term Matrix</i> .....	37
Tabel 4.4 Contoh Data tweet.....	40
Tabel 4.5 Ilustrasi Perhitungan NBC.....	41
Tabel 4.6 Ilustrasi Perhitungan NBC.....	42
Tabel 4.7 Akurasi <i>10-fold Cross Validation</i> .....	44
Tabel 4.8 <i>Confusion Matrix</i> pada <i>fold-5</i> .....	45
Table 4.9 Perhitungan Ketepatan Klasifikasi <i>Naïve Bayes</i> .....	45

## DAFTAR GAMBAR

Gambar 1.1 Data Pengadaan Vaksin lokasi Indonesia.....	2
Gambar 2.1 Ilustrasi Pembaggian Data .....	17
Gambar 3.1 Diagram Alir Praproses Teks .....	30
Gambar 3.2 Diagram Alir Klasifikasi <i>Naïve Bayes</i> .....	31
Gambar 4.1 Perbandingan Sentimen Positif dan Sentimen Negatif.....	36
Gambar 4.2 Frekuensi kemunculan kata.....	38
Gambar 4.3 Frekuensi kata pada Sentimen Positif.....	39
Gambar 4.4 Frekuensi kata pada Sentimen Negatif.....	39
Gambar 4.5 <i>Wordcloud</i> Data Tweet Vaksin Covid-19 .....	46
Gambar 4.6 <i>Wordcloud</i> Data Tweet Sentimen Positif.....	47
Gambar 4.7 <i>Wordcloud</i> Data Tweet Sentimen Negatif .....	47

## DAFTAR LAMPIRAN

Lampiran 1 Tahap Pengambilan <i>API Twitter</i> .....	52
Lampiran 1 Tahap Pengambilan <i>API Twitter</i> (Lanjutan).....	53
Lampiran 1 Tahap Pengambilan <i>API Twitter</i> (Lanjutan).....	54
Lampiran 1 Tahap Pengambilan <i>API Twitter</i> (Lanjutan).....	55
Lampiran 2 Proses Pengambilan Data.....	56
Lampiran 3 Praproses teks menggunakan <i>Rstudio</i> .....	56
Lampiran 3 Praproses teks menggunakan <i>Rstudio</i> (Lanjutan) .....	57
Lampiran 3 Praproses teks menggunakan <i>Rstudio</i> (Lanjutan) .....	58
Lampiran 4 Data Vaksin Covid (setelah praproses teks) .....	59
Lampiran 5 <i>Syntax</i> untuk analisis dengan <i>Google Colaboratory</i> .....	60
Lampiran 6 <i>Syntax NBC</i> menggunakan <i>Google Colaboratory</i> .....	62
Lampiran 7 Output Performa Klasifikasi Metode NBC.....	63