

**ANALISIS SENTIMEN KATA VAKSIN COVID-19 PADA TWITTER  
MENGUNAKAN METODE NAÏVE BAYES**

**SKRIPSI**



disusun oleh

**Rafly Inggar Deansyach**

**18.11.2017**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2022**

**ANALISIS SENTIMEN KATA VAKSIN COVID-19 PADA TWITTER  
MENGUNAKAN METODE NAÏVE BAYES**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Sistem Informasi



disusun oleh

**Rafly Inggar Deansyach**

**18.11.2017**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2022**

# PERSETUJUAN

## SKRIPSI

### ANALISIS SENTIMEN KATA VAKSIN COVID-19 PADA TWITTER MENGUNAKAN METODE NAÏVE BAYES

yang dipersiapkan dan disusun oleh

**Rafly Inggar Deansyach**

18.11.2017

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 10 Februari 2022

**Dosen Pembimbing,**

**Bayu Setiaji, M.Kom.**  
NIK. 190302216

**PENGESAHAN**  
**SKRIPSI**  
**ANALISIS SENTIMEN KATA VAKSIN COVID-19 PADA TWITTER**  
**MENGGUNAKAN METODE NAÏVE BAYES**

yang dipersiapkan dan disusun oleh

**Rafly Ingar Deansyach**

**18.11.2017**

telah dipertahankan di depan Dewan Penguji  
pada tanggal **17 Februari 2022**

**Susunan Dewan Penguji**

**Nama Penguji**

**Anggit Dwi Hartanto, M.Kom**  
**NIK. 190302163**

**Dina Maulina, M.Kom**  
**NIK. 190302250**

**Bayu Setiaji, M.Kom**  
**NIK. 190302216**

**Tanda Tangan**

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal **17 Februari 2022**

**DEKAN UNIVERSITAS AMIKOM YOGYAKARTA**

**Hanif Al Fatta, S.Kom, M. Kom**  
**NIK. 190302096**

## PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 19 Maret 2022



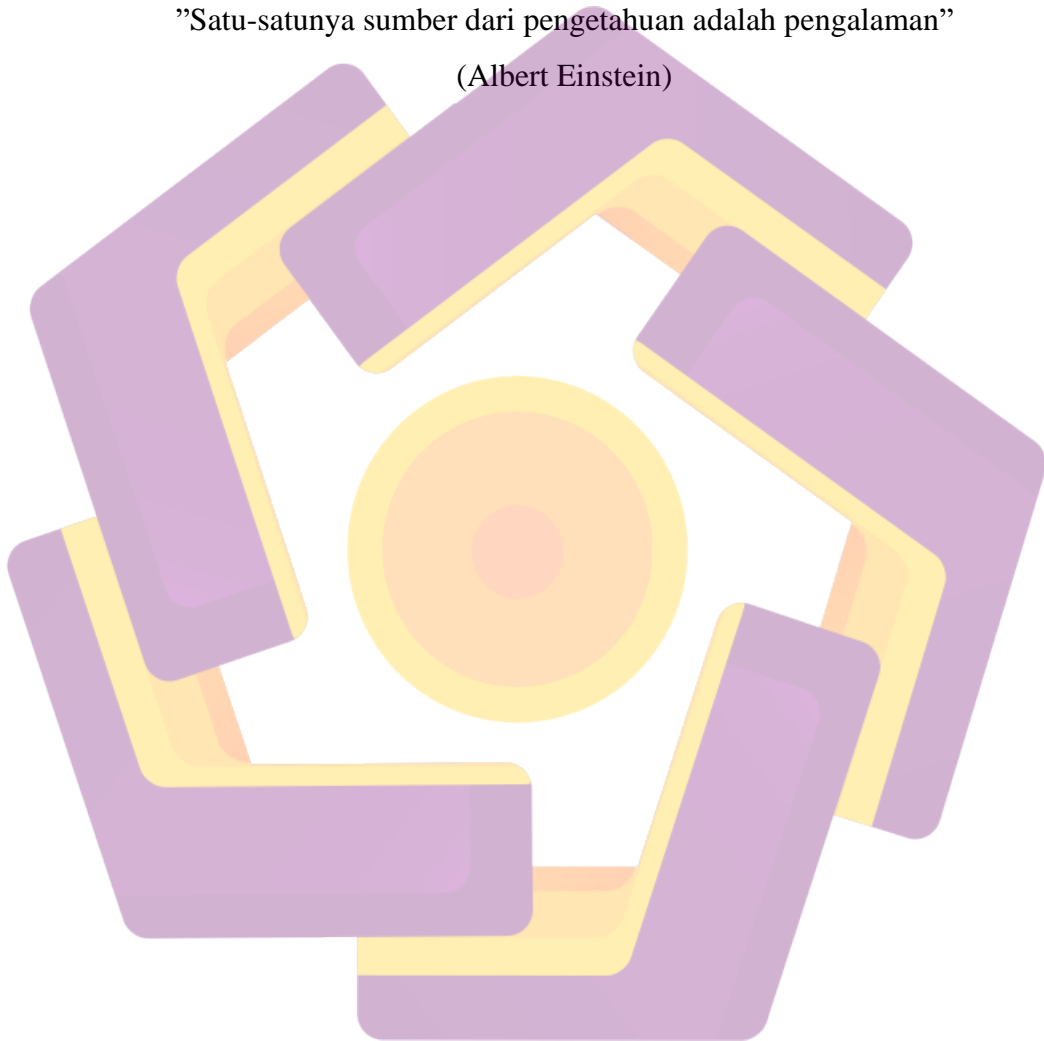
Rafly Inggar Deansyach  
NIM. 18.11.2017

## MOTTO

Tidak ada sebuah kesempatan yang datang secara kebetulan, karena adanya kesempatan selalu dibarengi dengan adanya usaha

”Satu-satunya sumber dari pengetahuan adalah pengalaman”

(Albert Einstein)



## PERSEMBAHAN

Puji syukur atas kehadiran Allah SWT yang senantiasa memberikan rahmatnya, taufik serta hidayah-Nya sehingga penulis diberi kemudahan dan kelancaran dalam menyelesaikan skripsi ini sampai akhir. Dengan ini penulis mempersembahkan skripsi ini kepada semua pihak yang mendukung kegiatan baik secara langsung maupun tidak langsung, yaitu:

1. Allah SWT yang selalu memberikan rahmat, nikmat dan karunia -Nya kepada penulis.
2. Kepada kedua orang tua, kakak, dan adik yang selalu mendoakan, memberi motivasi dan semangat sehingga penulis dapat menyelesaikan skripsi dengan baik
3. Keluarga besar yang sudah membantu dalam proses pengerjaan, memberikan dukungan serta selalu mendoakan.
4. Dosen pembimbing, Bayu Setiaji, M.Kom yang telah membimbing dan membantu penulis dari awal sampai akhir pembuatan skripsi.
5. Seluruh dosen dan staff Universitas Amikom Yogyakarta yang telah membantu semasa perkuliahan dengan ilmu dan fasilitasnya.
6. Teman – teman yang selalu membantu penulis dalam menyelesaikan permasalahan yaitu, Arvi Pramudyantoro, dan Laila Nur Aini.
7. Teman – teman semasa SMA yang selalu mendoakan, berjuang bersama, dan menghibur penulis.
8. Dan untuk semuanya yang tidak bisa penulis sebutkan satu persatu.

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah yang telah melimpahkan rahmat serta hidayah-Nya, tidak lupa shalawat serta salam selalu penulis panjatkan kepada junjungan kita Nabi Muhammad SAW, yang telah menuntun umatnya sehingga penulis dapat menyelesaikan skripsi dengan baik.

Skripsi yang berjudul “Analisis Sentimen Kata Vaksin COVID-19 pada Twitter menggunakan Metode Naive Bayes” disusun sebagai salah satu syarat utama untuk menyelesaikan program sarjana pada Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.

Pengerjaan skripsi ini tidak lepas dari bantuan berbagai pihak. Oleh Karena itu, penulis ingin menyampaikan rasa hormat dan terimakasih kepada:

1. Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Hanif Al Fatta, S.Kom, M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Windha Mega Pradnya D, M.Kom selaku Ketua Program Studi Informatika Universitas AMIKOM Yogyakarta.
4. Bapak Bayu Setiaji, M. Kom selaku dosen wali serta dosen pembimbing yang telah membantu penulis dengan saran dan waktunya.

Penulis menyadari sepenuhnya penelitian ini masih terdapat kekurangan, maka dari itu kritik dan saran serta masukan dari berbagai pihak akan penulis terima dengan lapang dada sebagai perbaikan karya – karya selanjutnya. Semoga skripsi yang sederhana ini dapat bermanfaat bagi pihak yang membutuhkan

Yogyakarta,

Rafly Inggar Deansyach



## DAFTAR ISI

<b>JUDUL</b> .....	II
<b>PERSETUJUAN</b> .....	III
<b>PENGESAHAN</b> .....	IV
<b>PERNYATAAN</b> .....	<b>ERROR! BOOKMARK NOT DEFINED.</b>
<b>MOTTO</b> .....	VII
<b>PERSEMBAHAN</b> .....	VIII
<b>KATA PENGANTAR</b> .....	VIII
<b>DAFTAR ISI</b> .....	IX
<b>DAFTAR TABEL</b> .....	XII
<b>DAFTAR GAMBAR</b> .....	XIII
<b>INTISARI</b> .....	XV
<b>ABSTRACT</b> .....	XVI
<b>BAB I PENDAHULUAN</b> .....	1
1.1 LATAR BELAKANG .....	1
1.2 RUMUSAN MASALAH .....	3
1.3 BATASAN MASALAH .....	3
1.4 MAKSUD DAN TUJUAN PENELITIAN .....	3
1.5 MANFAAT PENELITIAN .....	4
1.6 METODE PENELITIAN .....	4
1.6.1 Metode Pengumpulan Data.....	4
1.6.2 Metode Analisis .....	4
1.6.3 Metode Perancangan .....	5
1.6.4 Metode Implementasi .....	5
1.6.5 Metode Pegujian .....	5
1.7 SISTEMATIKA PENELITIAN.....	5
<b>BAB II LANDASAN TEORI</b> .....	7
2.1 TINJAUAN PUSTAKA .....	7
2.2 DASAR TEORI.....	9
2.2.1 Analisis Sentimen .....	10
2.2.2 COVID-19 .....	11

2.2.3	<i>Vaksin</i>	11
2.2.4	<i>Twitter</i>	12
2.2.5	<i>API Twitter</i>	12
2.2.6	<i>Klasifikasi</i>	12
2.2.7	<i>Naïve Bayes</i>	13
2.2.8	<i>Term Weigthing</i>	15
2.2.9	<i>Text Preprocessing</i>	15
2.2.10	<i>Confusion Matrix</i>	16
<b>BAB III ANALISIS DAN PERANCANGAN</b>		18
3.1	<b>ANALISIS KEBUTUHAN</b>	18
3.1.1	<i>Analisis Kebutuhan Fungsional</i>	18
3.1.2	<i>Analisis Kebutuhan Non Fungsional</i>	18
3.2	<b>FLOWCHART PROSES</b>	19
3.3	<b>PERANCANGAN SISTEM</b>	20
3.3.1	Pengambilan Data	20
3.3.2	<i>Pelabelan</i>	20
3.3.3	<i>Praproses Kata</i>	25
3.3.4	<i>Pembobotan Kata</i>	26
3.3.5	<i>Perhitungan Klasifikasi</i>	30
3.3.6	<i>Pengujian</i>	37
<b>BAB IV IMPLEMETASI DAN PEMBAHASAN</b>		40
4.1	<b>DESKRIPSI IMPLEMNTASI</b>	40
4.2	<b>IMPLEMENTASI SISTEM ANALISIS</b>	40
4.2.1	Pengambilan Data	40
4.2.2	<i>Pelabelan</i>	43
4.2.3	<i>Praproses Kata</i>	46
4.2.4	<i>Pembobotan Kata</i>	52
4.2.5	<i>Perhitungan Klasifikasi</i>	53
4.2.6	<i>Pengujian</i>	53
<b>BAB V PENUTUP</b>		57
5.1	<b>KESIMPULAN</b>	57

5.2 SARAN.....	57
<b>DAFTAR PUSTAKA.....</b>	<b>58</b>



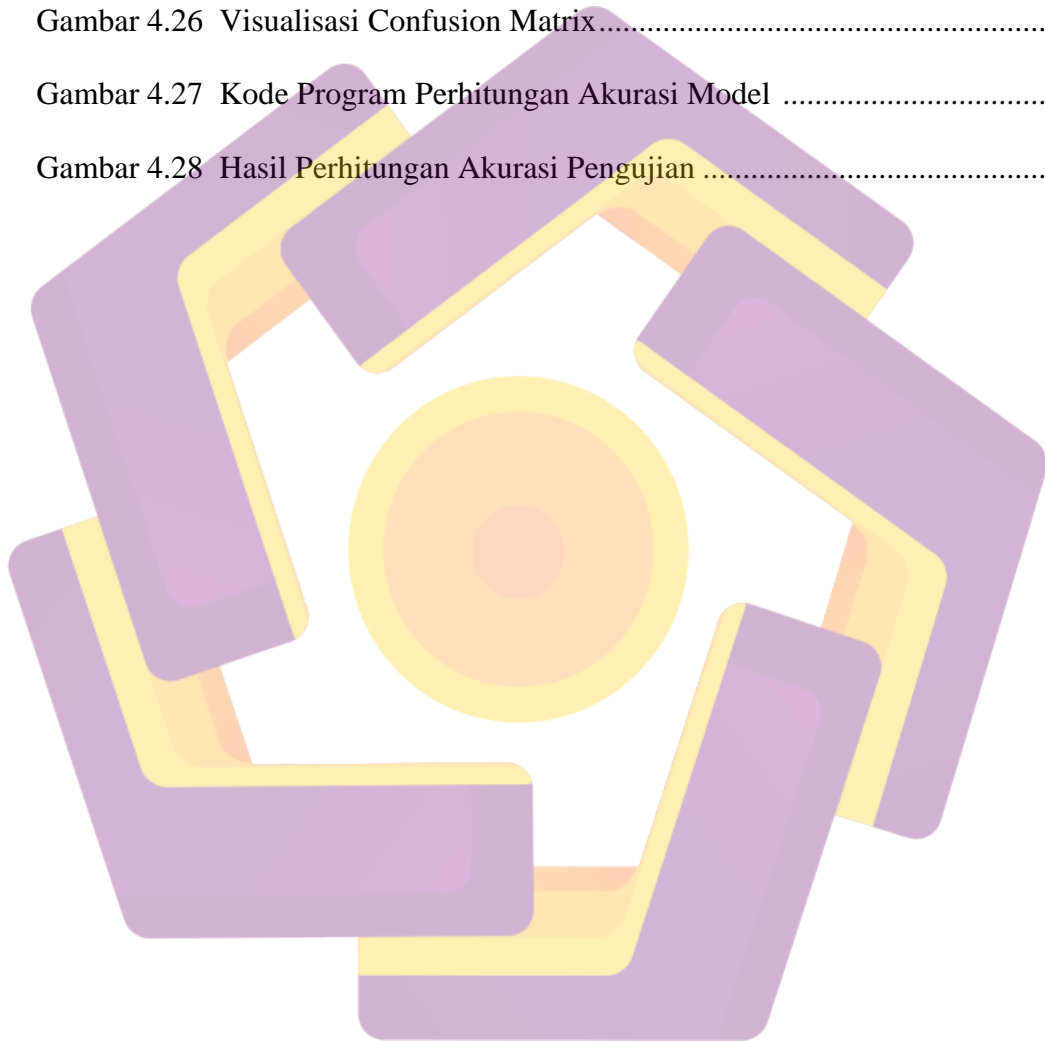
## DAFTAR TABEL

Tabel 2.1 Perbandingan penelitian.....	8
Tabel 2.2 Confusion Matrix .....	16
Tabel 3.1 Mengandung Sentimen Positif dan Negatif .....	21
Tabel 3.2 Contoh Proses Pelabelan.....	22
Tabel 3.3 Contoh Proses Cleaning .....	23
Tabel 3.4 Contoh Proses Case Folding .....	24
Tabel 3.5 Contoh Proses Tokenizing .....	24
Tabel 3.6 Contoh Proses Stopword Removal .....	25
Tabel 3.7 Contoh Proses Stemming .....	26
Tabel 3.8 Data Tweet .....	27
Tabel 3.9 Kumpulan data .....	28
Tabel 3.10 Perhitungan Term frequency (TF) .....	28
Tabel 3.11 Perhitungan Invers Document Frequency (IDF) .....	29
Tabel 3.12 Perhitungan (TF-IDF) .....	30
Tabel 3.13 Kalimat Data Latih.....	32
Tabel 3.14 Perhitungan Data Latih .....	33
Tabel 3.15 Data Uji .....	34
Tabel 3.16 Perhitungan Probabilitas Data Uji.....	37
Tabel 3.17 Perbandingan Data Uji dan Data Latih .....	38
Tabel 3.18 Perhitungan Manual Confusion Matrix.....	39
Tabel 4.1 Pembagian Data Latih dan Data Uji .....	52
Tabel 4.2 Hasil Pengujian Confusion Matrix.....	55

## DAFTAR GAMBAR

Gambar 3.1	Flowchart Proses .....	19
Gambar 4.1	Import Library .....	42
Gambar 4.2	Mendaftar Twitter Developer.....	42
Gambar 4.3	Token API Twitter .....	43
Gambar 4.4	Kode Program Crawling.....	43
Gambar 4.5	Tampilan Dataset .....	44
Gambar 4.6	Hasil Pelabelan Dataset.....	45
Gambar 4.7	Diagram Pelabelan Dataset .....	46
Gambar 4.8	Import Library .....	47
Gambar 4.9	Kode Program Cleaning .....	47
Gambar 4.10	Hasil Kode Program Cleaning.....	47
Gambar 4.11	Kode Program Case Folding .....	48
Gambar 4.12	Hasil Kode Program Case Folding.....	48
Gambar 4.13	Kode Program Tokenizing .....	49
Gambar 4.14	Hasil Kode Program Tokenixing.....	49
Gambar 4.15	Perintah Instal Library Sastrawi.....	50
Gambar 4.16	Import Library Sastrawi .....	50
Gambar 4.17	Kode Program Stopword.....	50
Gambar 4.18	Hasil Kode Program Stopword .....	50
Gambar 4.19	Import Library .....	51
Gambar 4.20	Kode Program Stemming.....	51
Gambar 4.21	Hasil Kode Program Stemming.....	51

Gambar 4.22 Kode Pembagian Data Latih dan Data Uji .....	52
Gambar 4.23 Kode Program TF-IDF dan Hasil Pembagian Data .....	53
Gambar 4.24 Kode Program Klasifikasi Naïve Bayes.....	53
Gambar 4.25 Kode Program Confusion Matrix.....	54
Gambar 4.26 Visualisasi Confusion Matrix.....	54
Gambar 4.27 Kode Program Perhitungan Akurasi Model .....	55
Gambar 4.28 Hasil Perhitungan Akurasi Pengujian .....	56



## INTISARI

Vaksin COVID-19 merupakan kebijakan pemerintah dalam rangka mencegah penyebaran virus COVID-19 yang terjadi di Indonesia saat ini, vaksin ini berguna untuk meningkatkan kekebalan tubuh seseorang terhadap suatu penyakit agar ketika terpapar virus tersebut akan tidak sakit atau hanya mengalami sakit ringan. Namun reaksi masyarakat terhadap kebijakan pemerintah ini menimbulkan tanggapan beragam maka Penelitian ini bertujuan untuk membuat model analisis sentimen terhadap opini masyarakat mengenai hal ini.

Analisis sentimen sendiri merupakan penambangan kontekstual berupa teks yang prosesnya mengidentifikasi dan memperoleh data dalam bentuk informasi subjektif teks yang didapat pada materi sumber dan hal ini membantu untuk memahami sentimen masyarakat, opini, penilaian seseorang, sikap dan emosi seseorang. Oleh Karena itu analisis sentimen digunakan untuk memilih dari suatu dataset pada bentuk teks dengan cara mendeteksi dan menyaringnya supaya tidak terjadi penyebaran informasi yang tidak benar, dengan menggunakan metode klasifikasi naïve bayes akan diketahui sentimen bernilai positif, negatif dan netral.

Hasil penelitian dengan menggunakan metode klasifikasi *Naïve Bayes* dengan pembobotan kata *tf-idf*, dari hasil pengujian analisis sentimen yang dapat dilakukan oleh sistem, mendapatkan *accuracy* adalah sebesar 75% dengan nilai rincian *precision* 76% dan *recall* 76%.

**Kata kunci:** Vaksin, COVID-19, Sentimen, *Naïve Bayes*, Klasifikasi Kata

## **ABSTRACT**

*The COVID-19 vaccine is a government policy in order to prevent the spread of the COVID-19 virus that is currently occurring in Indonesia, this vaccine is useful for increasing a person's immunity against a disease so that when exposed to the virus they will not get sick or only experience mild illness. However, the public's reaction to this government policy has given rise to various responses, so this study aims to create a sentiment analysis model for public opinion on this matter.*

*Sentiment analysis itself is a contextual mining in the form of text, the process of which is to identify and obtain data in the form of subjective text information obtained from the source material and this helps to understand people's sentiments, opinions, judgments, attitudes and emotions of a person. Therefore, sentiment analysis is used to select from a dataset in text form by detecting and filtering it so that there is no spread of incorrect information, by using the naive Bayes classification method it will be known that sentiment is positive, negative and neutral.*

*The results of the study using the Naïve Bayes classification method with tf-idf word weighting, from the results of sentiment analysis testing that can be carried out by the system, getting accuracy is 75% with 76% precision detail value and 76% recall.*

**Keyword:** Vaccine, COVID-19, Sentiment, Naïve Bayes, Word Classification