

**SENTIMEN ANALISIS PENGGUNA TWITTER
TERHADAP KEPUTUSAN PEMERINTAH
TERKAIT COVID-19 DENGAN
ALGORITMA NAÏVE BAYES**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Prodi Informatika



disusun oleh

Ikhwan Setiawan

18.11.1911

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2024**

**SENTIMEN ANALISIS PENGGUNA TWITTER
TERHADAP KEPUTUSAN PEMERINTAH
TERKAIT COVID-19 DENGAN
ALGORITMA NAÏVE BAYES**

SKRIPSI

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



disusun oleh

Ikhwan Setiawan

18.11.1911

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2024**

HALAMAN PERSETUJUAN

SKRIPSI

SENTIMEN ANALISIS PENGGUNA TWITTER TERHADAP KEPUTUSAN PEMERINTAH TERKAIT COVID-19 DENGAN ALGORITMA NAÏVE BAYES

yang disusun dan diajukan oleh

Ikhwan Setiawan

18.11.1911

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 22 Mei 2023

Dosen Pembimbing,


Ainal Yaqin.M.Kom
NIK. 190302255

HALAMAN PENGESAHAN

SKRIPSI

SENTIMEN ANALISIS PENGGUNA TWITTER TERHADAP KEPUTUSAN PEMERINTAH TERKAIT COVID-19 DENGAN ALGORITMA NAÏVE BAYES

yang disusun dan diajukan oleh

Ikhwan Setiawan

18.11.1911

Telah dipertahankan di depan Dewan Penguji
pada tanggal 22 Mei 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ainul Yaqin, M.Kom
NIK. 190302255

Arifiyanto Hadinegoro, S.Kom, MT
NIK. 190302289

Anggit Ferdita Nugraha, S.T., M.Eng
NIK. 190302480

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 22 Mei 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Ikhwan Setiawan
NIM : 18.11.1911

Menyatakan bahwa Skripsi dengan judul berikut:

**Sentimen Analisis Pengguna Twitter Terhadap Keputusan Pemerintah
Terkait Covid-19 Dengan Algoritma Naïve Bayes**

Dosen Pembimbing : Ainul Yaqin, M.Kom

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 22 Mei 2023

Yang Menyatakan,



Ikhwan Setiawan

HALAMAN PERSEMBAHAN

Puji syukur kami panjatkan ke hadirat Allah SWT yang telah memberikan rahmat dan hidayah-Nya serta kasih sayang-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi ini,

Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan program studi S1 di Universitas Amkomu Yogyakarta. Dalam proses penyusunan skripsi ini, penulis mengucapkan terima kasih kepada:

1. Orang Tua dan Keluarga, yang selalu memberikan dukungan moral dan material dalam proses pendidikan penulis hingga saat ini.
2. Bapak Ainul Yaqin M.Kom, sebagai dosen pembimbing yang telah memberikan bimbingan, arahan, dan motivasi selama proses penyusunan skripsi ini.
3. Teman-teman dan Rekan-rekan, yang telah memberikan dukungan dan bantuan selama proses penyusunan skripsi ini.

Semoga hasil dari penulisan skripsi ini dapat bermanfaat bagi pembaca dan dapat menjadi sumber inspirasi bagi peneliti selanjutnya dalam perkembangan teknologi dan informasi di masa yang akan datang.

Kata Pengantar

Dengan segala rasa syukur, puji dan syukur kehadiran Allah SWT Yang Maha Esa, penulis mengucapkan terima kasih yang setinggi-tingginya kepada semua pihak yang telah membantu dalam menyelesaikan skripsi ini. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana program studi Sistem Informasi pada Universitas Amikom Yogyakarta.

Penulisan skripsi ini tentunya tidak bisa terlaksana dengan baik tanpa bantuan dan dukungan dari banyak pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada dosen pembimbing bapak Ainul Yaqin M.Kom yang telah memberikan arahan, saran, dan bimbingan yang sangat berharga dalam penyusunan skripsi ini. Terima kasih juga kepada para dosen pengajar di program studi Sistem Informasi, yang telah memberikan ilmu dan pengetahuan yang bermanfaat selama perkuliahan.

Penulisan skripsi ini tentunya memiliki banyak kekurangan dan kesalahan, oleh karena itu penulis memohon maaf yang sebesar-besarnya. Semoga skripsi ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, pengembangan teknologi dan pengembangan keilmuan di masa yang akan datang terkhusus dalam dunia pendidikan.

Yogyakarta, 22 Mei 2023

Ikhwan Setiawan

DAFTAR ISI

SENTIMEN ANALISIS PENGGUNA TWITTER TERHADAP

KEPUTUSAN PEMERINTAH TERKAIT COVID-19

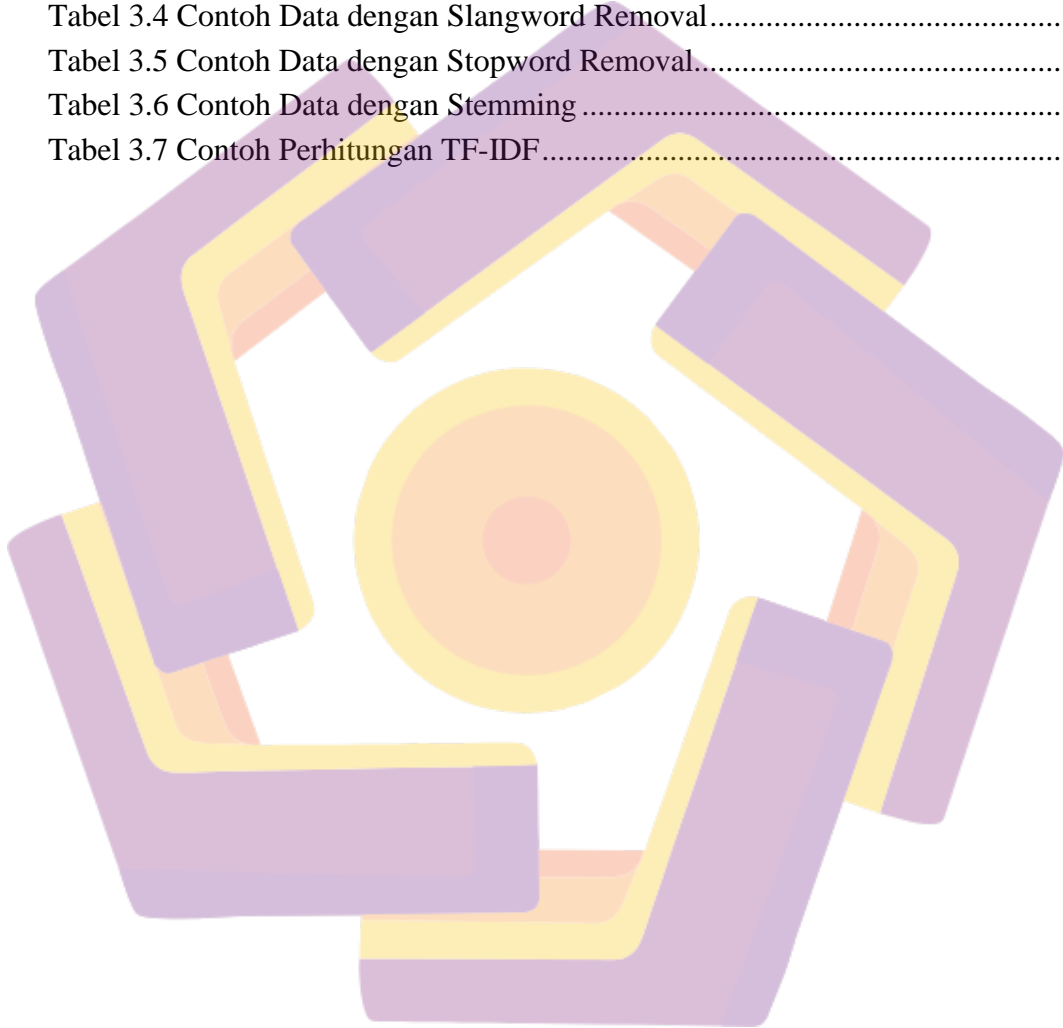
DENGAN ALGORITMA NAÏVE BAYES	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR RUMUS	xii
INTISARI	xiii
ABSTRACT.....	ix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Maksud dan Tujuan Penelitian	3
1.5 Manfaat Penelitian.....	3

1.6	Metode Penelitian.....	3
1.6.1	Metode Pengumpulan Data	3
1.6.2	Metode Analisis.....	4
1.7	Sistematika Penulisan.....	4
BAB II LANDASAN TEORI		6
2.1	Kajian Pustaka.....	6
2.1	Landasan Teori.....	11
2.1.1	Analisis Sentimen.....	11
2.1.2	Data Mining	11
2.1.3	Twitter.....	11
2.1.4	Text Preprocessing	11
2.1.5	Natural Language Processing.....	13
2.1.7	Naïve Bayes	14
2.1.8	Confusion Matrix.....	16
BAB III METODOLOGI PENELITIAN.....		18
3.1	Objek Penelitian.....	18
3.2	Alur Penelitian	18
3.3	Alat dan Bahan.....	20
3.3.1	Kebutuhan Fungsional.....	20
3.3.2	Kebutuhan Non-Fungsional	20
3.4	Pengumpulan Data	21
3.5	Cleaning Data.....	21
3.6	Case Folding.....	23
3.7	Slangword Removal	25
3.8	Stopword Removal.....	28

3.9	Stemming	29
3.10	Term Frequency – Invers Document Frequency (TF-IDF).....	31
3.11	Algoritma Naïve Bayes	34
3.12	Precision, Recall dan F-Measure.....	36
BAB IV IMPLEMENTASI DAN PEMBAHASAN.....		37
4.1	Deskripsi Implementasi.....	37
4.2	Implementasi Pengumpulan Data Tweet.....	37
4.3	Implementasi Pelabelan Data Training	39
4.4	Implementasi Perancangan Fungsi dan Sistem.....	39
4.5	Pembuatan Fungsi Pembersih data (Cleaning).....	40
4.4.1	Pembuatan Fungsi Stopwords / Stop Removal	41
4.4.2	Pembuatan Fungsi Stemming.....	42
4.4.3	Pembuatan Fungsi Text Vectorization	44
4.4.4	Pembuatan Fungsi Klasifikasi Naïve Bayes.....	44
4.4.5	Evaluasi Model Klasifikasi	45
BAB V KESIMPULAN.....		47
5.1	Kesimpulan.....	47
5.2	Saran.....	47
DAFTAR PUSTAKA		48

DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka.....	6
Tabel 2.1 Confusion Matrix.....	11
Tabel 3.1 Contoh Data Cleaning.....	14
Tabel 3.2 Contoh Case Folding	16
Tabel 3.3 Contoh Data kata tidak baku.....	17
Tabel 3.4 Contoh Data dengan Slangword Removal.....	18
Tabel 3.5 Contoh Data dengan Stopword Removal.....	20
Tabel 3.6 Contoh Data dengan Stemming	21
Tabel 3.7 Contoh Perhitungan TF-IDF.....	23



DAFTAR GAMBAR

Gambar 3.1 Diagram Alir Proses Klasifikasi Naïve Bayes.....	26
Gambar 3.2 Konfigurasi Crawling Data	27
Gambar 3.3 File Training.....	28
Gambar 4.1 Script crawling data twitter.....	30
Gambar 4.2 Hasil Crawling data.....	30
Gambar 4.3 Dataset yang sudah berlabel.....	31
Gambar 4.5 Hasil penghapusan tweet yang duplikasi	32
Gambar 4.6 Script Cleaning Data.....	32
Gambar 4.7 Hasil Cleaning Data	33
Gambar 4.8 Script Stopwords	33
Gambar 4.9 Hasil Stopwords	33
Gambar 4.10 Script Stemming.....	34
Gambar 4.11 Hasil Stemming.....	34
Gambar 4.12 Script TF-IDF.....	35
Gambar 4.13 merupakan script klasifikasi naïve bayes.....	35
Gambar 4.14 Visualisasi Confusion Matrix	38

DAFTAR RUMUS

Rumus 2.1 Pembobotan Kata TF-IDF	8
Rumus 2.2 Persamaan Teorema Naïve Bayes	9
Rumus 2.3 Persamaan Confusion Matrix	12
Rumus 3.1 Persamaan TF-IDF	26
Rumus 3.2 Persamaan probabilitas prior	29
Rumus 3.3 Persamaan probabilitas prior	29
Rumus 3.4 Probabilitas kata ke-n	29
Rumus 3.5 Menghitung Recall dan F-Measure.....	30
Rumus 4.1 Pengujian Precision positif.....	40
Rumus 4.2 Pengujian Recall positif	40
Rumus 4.3 Pengujian F-Measure positif	40
Rumus 4.4 Pengujian Precision negatif	40
Rumus 4.5 Pengujian Recall negatif.....	40
Rumus 4.6 Pengujian F-Measure negatif	40
Rumus 4.7 Pengujian Precision netral	40
Rumus 4.8 Pengujian Recall netral.....	40
Rumus 4.9 Pengujian F-Measure netral	40
Rumus 4.10 Persamaan Nilai akurasi	41

INTISARI

Aspirasi, opini dan kritik masyarakat sangat sulit disampaikan secara langsung, karena ada beberapa kendala seperti keterbatasan waktu, ruang, bahkan SDM. Namun, belakangan ini jejaring sosial tidak menutup kemungkinan akan hal tersebut, salah satunya adalah dengan twitter. Twitter merupakan salah satu jejaring sosial yang sering digunakan oleh masyarakat untuk mengungkapkan opini, kritik hingga aspirasi. Namun opini pada twitter banyak sekali makna yang berbeda dari setiap netizen, salah satunya adalah tentang opini COVID-19 terhadap Pemerintah. Oleh karena itu perlu suatu analisis sentimen opini masyarakat guna menyalurkan dan memberi pandangan baru mengenai suatu isu tentang COVID-19 yang menimbulkan banyak kerancuan dimasyarakat, metode yang digunakan dalam setimen analisis ini adalah algoritma Naïve Bayes, penelitian ini bertujuan mengetahui kecenderungan opini masyarakat di twitter. Subjek diambil menggunakan API twitter sebanyak dengan menggunakan kata kunci “COVID-19” dan “Pemerintah”, Dimulai dari klasifikasi opini positif atau negatif kemudian dilanjutkan dengan data cleansing dan Langkah terakhir yaitu dengan melakukan preprocessing.

Kata Kunci: Sentimen Analisis, Twitter, Covid-19, Pemerintah, Rakyat, Naïve Bayes

ABSTRACT

People's aspirations, opinions and criticisms are very difficult to convey directly, because there are several obstacles such as limited time, space, and even human resources. However, lately social networking does not rule out the possibility of this, one of which is with twitter. Twitter is a social network that is often used by the public to express opinions, criticisms and aspirations. However, opinions on Twitter have a lot of different meanings from each netizen, one of which is about the opinion of COVID-19 against the Government. Therefore, an analysis of public opinion sentiment is needed to harmonize and provide new views on an issue about COVID-19 which causes a lot of confusion in the community, the method used in this analysis is the Naïve Bayes algorithm, this study aims to determine the tendency of public opinion on twitter. Subjects were taken using the Twitter API as much as using the keywords "COVID-19" and "Government", Starting from the classification of positive or negative opinions, then proceeding with data cleansing and the final step was to do is preprocessing.

Keyword: Sentiment Analysis, Twitter, Covid-19, Government, People, Naïve Bayes