

**ANALISIS SENTIMEN TERHADAP PENANGANAN COVID-19 OLEH
PEMERINTAH INDONESIA PADA TWITTER MENGGUNAKAN
ALGORITMA NAIVE BAYES CLASSIFIER**

SKRIPSI



disusun oleh

Denny Saryanto

17.11.1234

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**ANALISIS SENTIMEN TERHADAP PENANGANAN COVID-19 OLEH
PEMERINTAH INDONESIA PADA TWITTER MENGGUNAKAN
ALGORITMA NAIVE BAYES CLASSIFIER**

SKRIPSI

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



disusun oleh

Denny Saryanto

17.11.1234

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

ANALISIS SENTIMEN TERHADAP PENANGANAN COVID-19 OLEH PEMERINTAH INDONESIA PADA TWITTER MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER

yang dipersiapkan dan disusun oleh

Denny Saryanto

17.11.1234

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 27 April 2021

Dosen Pembimbing,

Ike Verawati, M.Kom.

NIK. 190302237

PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN TERHADAP PENANGANAN COVID-19 OLEH PEMERINTAH INDONESIA PADA TWITTER MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER

yang dipersiapkan dan disusun oleh

Denny Saryanto

17.11.1234

telah dipertahankan di depan Dewan Penguji
pada tanggal 21 April 2021

Susunan Dewan Penguji

Nama Penguji

Yuli Astuti, M.Kom.
NIK. 190302146

Sri Ngudi Wahyuni, S.T., M.Kom.
NIK. 190302060

Ike Verawati, M.Kom.
NIK. 190302237

Tanda Tangan

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 10 Mei 2021

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, M.Kom.

NIK. 190302096

PERNYATAAN

PERNYATAAN

Saya yang bertandatangan di bawah 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, 10 Mei 2021

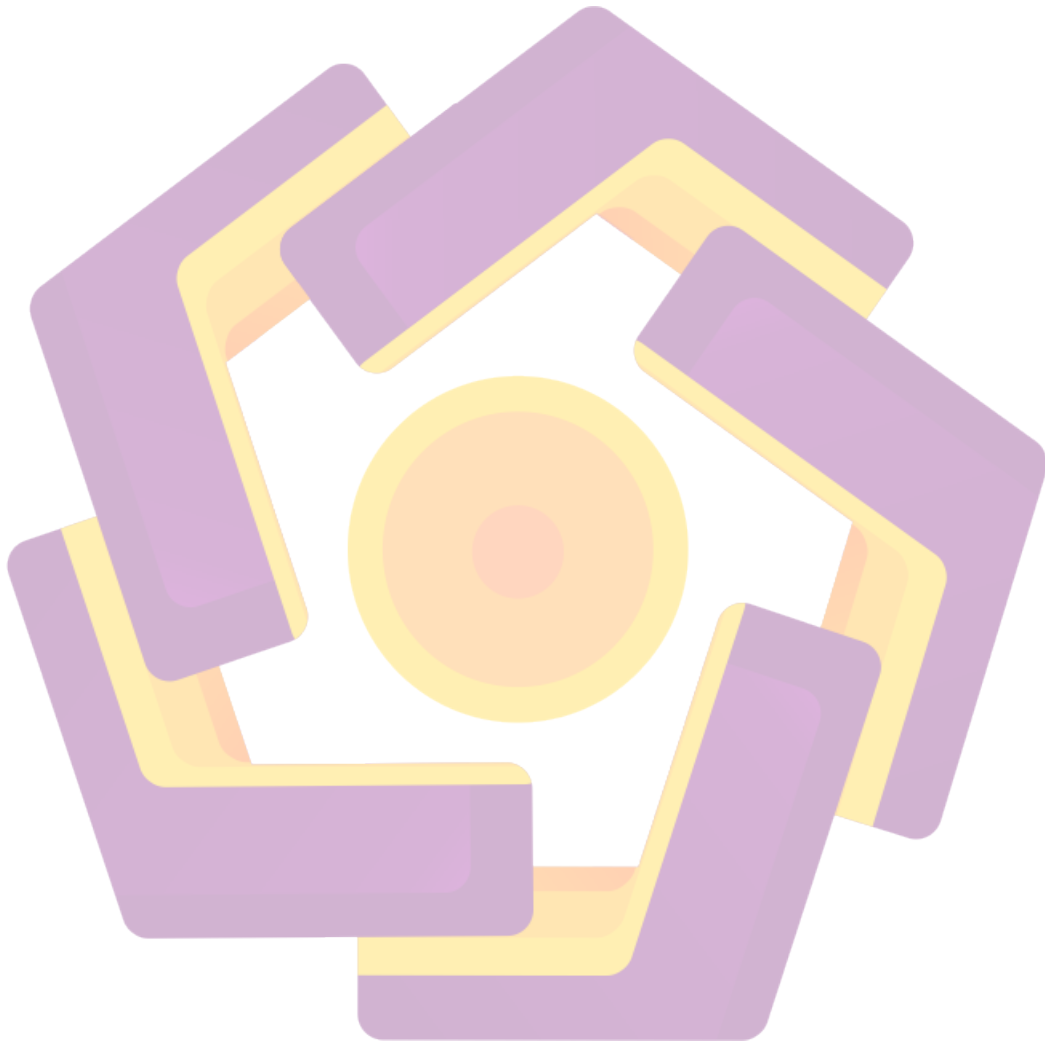


Denny Saryanto

NIM. 17.11.1234

MOTTO

“Berusaha Semaksimal Mungkin, Nikmati Prosesnya, Terima Apapun Hasilnya”



PERSEMBAHAN

Dengan rasa syukur yang dalam, saya mempersembahkan skripsi untuk berbagai pihak yang telah membantu dan mendukung peneliti untuk menyelesaikan skripsi.

1. Allah SWT yang selalu memberikan berkah, nikmat, dan kesehatan sehingga peneliti mampu menyelesaikan skripsi dengan baik.
2. Orang tua yang selalu mendukung, mengingatkan, dan memberi perhatian sepenuh hati agar peneliti selalu fokus dan serius dalam menyelesaikan skripsi.
3. Bu Ike Verawati selaku dosen pembimbing yang selalu memberikan saran dan membimbing dengan sangat baik serta mudah dimengerti dan dipahami.
4. Dosen penguji yang memberikan saran sehingga skripsi menjadi lebih baik lagi dan sesuai.
5. Dosen – dosen Universitas Amikom Yogyakarta yang telah memberikan ilmu yang sangat bermanfaat.
6. Teman – teman IF-05 yang telah menemani dari awal semester hingga selesai studi.

KATA PENGANTAR

Puji Syukur kepada Allah SWT yang senantiasa memberikan hidayah, nikmat, dan kesehatan kepada hamba-Nya. Sholawat serta salam tidak lupa kita curahkan kepada junjungan kita Nabi Muhammad SAW. Sehingga peneliti dapat menyelesaikan skripsi yang berjudul **“ANALISIS SENTIMEN TERHADAP PENANGANAN COVID-19 OLEH PEMERINTAH INDONESIA PADA TWITTER MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER”**.

Penelitian ini dibuat dan diajukan guna memenuhi dan menyelesaikan Program Studi Informatika di Universitas Amikom Yogyakarta. Peneliti menyadari penelitian ini masih jauh dari kata sempurna, baik dalam penulisan maupun materi dikarenakan keterbatasan pengetahuan peneliti dan kemampuan penulisan dari peneliti. Semoga dengan adanya penelitian ini dapat memberikan manfaat dan lebih baik lagi untuk penelitian selanjutnya.

Yogyakarta, 10 Mei 2021

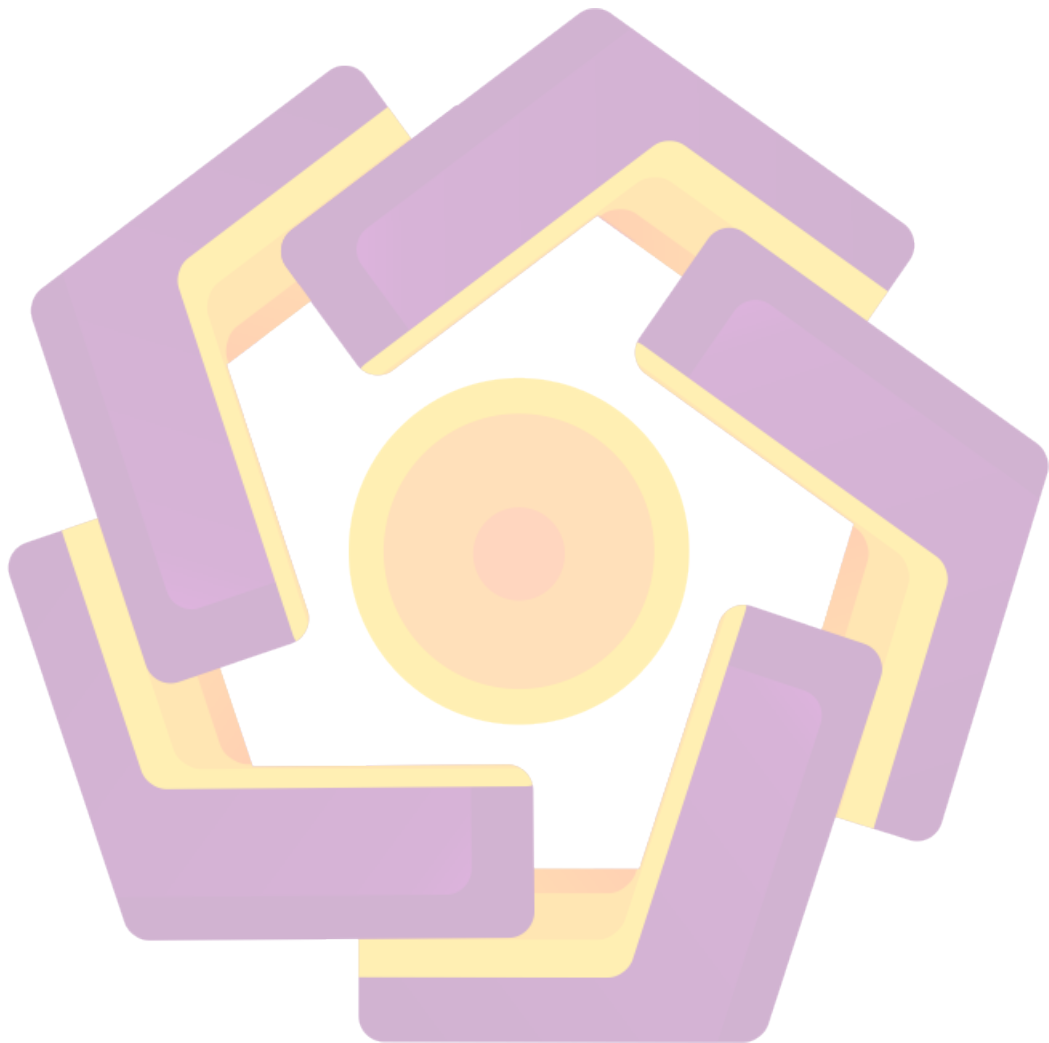
Denny Saryanto
17.11.1234

DAFTAR ISI

PERSETUJUAN	i
PENGESAHAN	ii
PERNYATAAN.....	iii
MOTTO	iv
PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
INTISARI.....	xiii
<i>ABSTRACT</i>	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Metode Penelitian.....	4
1.6.1 Metode Pengumpulan Data	4
1.6.2 Metode Analisis	5
1.6.3 Metode Implementasi.....	5
1.6.4 Metode Pengujian.....	6
1.7 Sistematika Penulisan.....	6

BAB II	LANDASAN TEORI.....	7
2.1	Tinjauan Pustaka	7
2.2	Python.....	10
2.3	Analisis Sentimen.....	10
2.4	<i>Text Mining</i>	10
2.5	<i>Lexicon Based</i>	11
2.6	<i>Text Preprocessing</i>	11
2.7	TF-IDF.....	12
2.8	Naïve Bayes.....	13
2.9	Metode Evaluasi.....	13
BAB III	METODOLOGI PENELITIAN	16
3.1	Metode Penelitian.....	16
3.2	Analisis Data	20
3.2.1	Spesifikasi Data.....	20
3.3	Analisis Kebutuhan	21
3.3.1	Analisis Kebutuhan Fungsional	21
3.3.2	Analisis Kebutuhan Non-Fungsional.....	22
3.4	Naïve Bayes Classifier	22
BAB IV	IMPLEMENTASI DAN PEMBAHASAN	23
4.1	<i>Scrape Data</i>	23
4.2	<i>Text Preprocessing</i>	23
4.2.1	Konversi <i>emoji</i> dan <i>emoticon</i>	23
4.2.2	<i>Text Cleaning</i>	24
4.2.3	Tokenisasi	25
4.2.4	Normalisasi teks	25

4.2.5	<i>Stemming</i>	26
4.2.6	<i>Stopwords Removal</i>	27
4.3	Pemberian Kelas	28
4.4	TF-IDF	28
4.5	Naïve Bayes Classifier	30
4.5.1	Multinomial Naïve Bayes	30
4.6	Confusion Matrix	34
4.7	Implementasi	36
4.7.1	Deklarasi Dataset	36
4.7.2	Konversi Emoji dan Emoticon	37
4.7.3	<i>Text Cleaning</i>	39
4.7.4	Tokenisasi	39
4.7.5	Normalisasi	40
4.7.6	<i>Stemming</i>	41
4.7.7	<i>Stopwords Removal</i>	41
4.7.8	<i>Lexicon Based</i>	43
4.7.9	TF-IDF	45
4.7.10	Data Latih dan Data Uji	46
4.7.11	Melatih Model	46
4.7.12	Hasil Pengujian	46
BAB V	PENUTUP	48
5.1	Kesimpulan	48
5.2	Saran	49
	DAFTAR PUSTAKA	50
	LAMPIRAN	52



DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka	8
Tabel 2.2 <i>Confusion Matrix</i>	14
Tabel 4.1 Hasil <i>Scrape</i> Tweet.	23
Tabel 4.2 Hasil Tweet Dengan Konversi Emoji	24
Tabel 4.3 Data Tweet Clean.....	24
Tabel 4.4 Tweet Tokenisasi.....	25
Tabel 4.5 Data Tweet Normalisasi.....	26
Tabel 4.6 Hasil Proses <i>Stemming</i>	26
Tabel 4.7 Data Tweet <i>Stopwords Removal</i>	27
Tabel 4.8 Sentimen Tweet.....	28
Tabel 4.9 Sampel Data.	28
Tabel 4.10 <i>Prior Probability</i>	31
Tabel 4.11 Contoh Data Uji	32
Tabel 4.12 Klasifikasi Data Uji.....	33
Tabel 4.13 Hasil Klasifikasi Data Uji	33
Tabel 4.14 Contoh Hasil Klasifikasi Model.....	34
Tabel 4.15 Hasil <i>Confusion Matrix</i>	34

DAFTAR GAMBAR

Gambar 4.1 Alur Penelitian.....	17
Gambar 4.2 Sampel Data Tweet	21
Gambar 4.3 Deklarasi Dataset.....	36
Gambar 4.4 Dataset <i>Emoji</i>	37
Gambar 4.5 Dataset <i>Emoticon</i>	38
Gambar 4.6 Kode Konversi Emoji dan Emoticon.	38
Gambar 4.7 Kode <i>Text Cleaning</i>	39
Gambar 4.8 Kode Tokenisasi.....	40
Gambar 4.9 Kode Normalisasi Dataset.....	40
Gambar 4.10 Dataset Normalisasi.....	40
Gambar 4.11 Kode <i>Stemming</i>	41
Gambar 4.12 Dataset <i>Stopwords</i>	42
Gambar 4.13 Kode <i>Remove Stopwords</i>	42
Gambar 4.14 <i>Lexicon</i> Negatif.	43
Gambar 4.15 <i>Lexicon</i> Positif.....	44
Gambar 4.16 Hasil Proses <i>Lexicon Based</i>	44
Gambar 4.17 Kumpulan <i>Feature List</i>	45
Gambar 4.18 Kode TF-IDF.....	45
Gambar 4.19 Data latih dan Data Uji.....	46
Gambar 4.20 Kode Melatih Model.	46
Gambar 4.21 <i>Confusion Matrix</i>	47
Gambar 4.22 Performa Model.....	47
Gambar 4.23 Hasil Klasifikasi 100 Data Tweet.....	47

INTISARI

Sosial media di masa sekarang memiliki peran yang sangat banyak bagi masyarakat dunia. Sosial media dapat dijadikan sebagai tempat berbagi bersama teman, berbagi informasi, diskusi suatu topik, hingga kepentingan strategi bisnis. Twitter adalah salah satu sosial media populer dan terdapat informasi – informasi yang dibagikan oleh pengguna Twitter. Informasi tersebut dapat berbentuk *tweet*, gambar, ataupun video.

Analisis teks meliputi *text preprocessing*, pemberian label kelas menggunakan *lexicon InSet*, merubah kalimat atau *tweet* menjadi vektor menggunakan TF-IDF, dan untuk menyelesaikan masalah klasifikasi menggunakan Naïve Bayes Classifier. Performa model diketahui dengan menggunakan *confusion matrix*.

Hasil klasifikasi model mendapatkan akurasi sebesar 56% dan dari 100 *tweet* yang dilakukan uji coba, menghasilkan 92% positif, 8% negatif, dan 0% netral. Dataset masih terdapat kekurangan dikarenakan, masih ada adanya data yang tidak diperlukan.

Kata Kunci : PSBB, Analisis Sentimen, Naïve Bayes, TF-IDF.

ABSTRACT

Today's social media has a very large role for the world community. Social media can be used as a place to share with friends, share information, discuss a topic, and even business strategy interests. Twitter is one of the popular social media and there is information that is shared by Twitter users. This information can be in the form of tweets, pictures or videos.

Text analysis includes text preprocessing, class labeling using InSet lexicon, converting sentences or tweets into vectors using TF-IDF, and solving classification problems using the Naïve Bayes Classifier. Model performance is known by using confusion matrix.

The classification results of the model get an accuracy of 56% and out of 100 tweets tested, the results are 92% positive, 8% negative, and 0% neutral. There are still deficiencies in the dataset because there is still unnecessary data.

Keywords : PSBB, Sentiment Analysis, Naïve Bayes, TF-IDF