

**ANALISIS SENTIMEN TWITTER TERHADAP TIKTOK SHOP DENGAN
ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE**



Vincent

00000041875

UMMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2023

**ANALISIS SENTIMEN TWITTER TERHADAP TIKTOK SHOP DENGAN
ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE**



Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Vincent

00000041875

UMMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2023

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Vincent
Nomor Induk Mahasiswa : 00000041875
Program Studi : Informatika

Skripsi dengan judul:

Analisis Sentimen Twitter Terhadap TikTok Shop Dengan Algoritma Naive Bayes dan Support Vector Machine

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan Skripsi maupun dalam penulisan laporan Skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas akhir yang telah saya tempuh.

Tangerang, 14 Juni 2023



(Vincent)

UMM
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISIS SENTIMEN TWITTER TERHADAP TIKTOK SHOP DENGAN ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE

oleh

Nama : Vincent
NIM : 00000041875
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Selasa, 21 Juni 2023

Pukul 15.00 s/s 17.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang



(Moeljono Widjaja, B.Sc., M.Sc., Ph.D)

NIDN: 0311106903

Penguji



(Farica Perdana Putri, S.Kom., M.Sc.)

NIDN: 0331019301

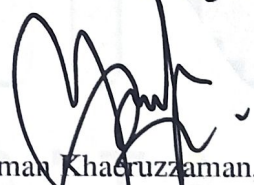
Pembimbing I



(Vincentius Kurniawan, S.Kom.,
M.Eng.Sc.)

NIDN: 0308079501

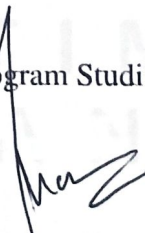
Pembimbing II



(Yaman Khaeruzzaman, M.Sc.)

NIDN: 0413057104

Ketua Program Studi Informatika,



(Marlinda Vasty Overbeek, S.Kom., M.Kom.)

NIDN: 0818038501

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Multimedia Nusantara, saya yang bertanda tangan di bawah ini:

Nama : Vincent
NIM : 00000041875
Program Studi : Informatika
Fakultas : Teknik dan Informatika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada **Universitas Multimedia Nusantara** hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

**ANALISIS SENTIMEN TWITTER TERHADAP TIKTOK SHOP DENGAN
ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Multimedia Nusantara berhak menyimpan, mengalih media / format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 14 Juni 2023

Yang menyatakan



Vincent

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Halaman Persembahan / Motto

”The fear of the LORD is the beginning of wisdom, And the knowledge
of the Holy One is understanding”

Proverbs 9:10 (NKJV)



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Analisis Sentimen Twitter Terhadap TikTok Shop Dengan Algoritma Naive Bayes dan Support Vector Machine dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Marlinda Vasty Overbeek, S.Kom., M.Kom., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Vincentius Kurniawan, S.Kom., M.Eng.Sc., sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
5. Yaman Khaeruzzaman, M.Sc., sebagai Pembimbing kedua yang telah banyak membantu dan memberikan bimbingan atas terselesainya Skripsi/Tesis ini.
6. Orang Tua, keluarga, dan teman saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tesis ini.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 14 Juni 2023



Vincent

ANALISIS SENTIMEN TWITTER TERHADAP TIKTOK SHOP DENGAN ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE

Vincent

ABSTRAK

Dengan perkembangan internet yang begitu pesat, istilah *e-commerce* sudah tidak asing lagi bagi masyarakat. Dalam beberapa tahun terakhir, *e-commerce* terus menerus mengalami pertumbuhan dengan semakin banyak orang yang beralih untuk berbelanja secara *online*. TikTok mulai memasuki *e-commerce* dengan meluncurkan TikTok shop yang merupakan fitur baru yang memungkinkan penggunanya untuk menelusuri dan membeli produk langsung di dalam aplikasi TikTok. Dengan semakin banyaknya *e-commerce* yang bermunculan maka diperlukan sebuah penilaian berdasarkan analisis sentimen pada sosial media Twitter yang diungkapkan dalam teks yang berupa opini positif atau negatif. Twitter memungkinkan penggunanya untuk saling berbagi pendapat atau opini melalui *tweet* yang dibatasi hingga 280 karakter. *Tweets* yang dibagikan ini dikumpulkan dan dijadikan data dalam penelitian ini. Pada penelitian ini, analisis sentimen dilakukan dengan menggunakan metode *Naive Bayes* dan *Support Vector Machine*. *Confusion matrix* kemudian digunakan untuk mendapatkan nilai *accuracy*, *precision*, *recall*, dan *f1-score*. Hasil dari penelitian ini menunjukkan metode *Support Vector Machine kernel linear* mendapatkan hasil lebih baik dibandingkan metode *Naive Bayes Multinomial* dengan perbandingan dataset 80:20 dengan hasil *accuracy* 83.36% untuk *Support Vector Machine Kernel Linear* sedangkan hasil *accuracy* 79.83% untuk *Naive Bayes Multinomial*.

Kata kunci: Analisis Sentimen, *Confusion Matrix*, *Naive Bayes*, *Support Vector Machine*, TikTok Shop, Twitter

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Sentiment Analysis of Twitter on TikTok Shop Using Naive Bayes and Support Vector Machine Algorithms

Vincent

ABSTRACT

With the rapid development of the internet, the term "e-commerce" has become familiar to the public. In recent years, e-commerce has experienced continuous growth, with an increasing number of people shifting towards online shopping. TikTok has entered the e-commerce space by launching TikTok Shop, a new feature that allows users to browse and purchase products directly within the TikTok application. As more e-commerce platforms emerge, there is a need for sentiment analysis based on Twitter social media, which contains positive or negative opinions. Twitter enables users to share their opinions or sentiments through tweets, limited to 280 characters. These shared tweets are collected and used as data in this study. Sentiment analysis is performed using the Naive Bayes and Support Vector Machine methods. The confusion matrix is then used to obtain accuracy, precision, recall, and f1-score values. The results of this research indicate that the method of Support Vector Machine with linear kernel performs better compared to the Naive Bayes Multinomial method with an 80:20 dataset split. The accuracy achieved is 83.36% for Support Vector Machine with linear kernel and 79.83% for Naive Bayes Multinomial.

Keywords: *Confusion Matrix, Naive Bayes, Sentiment Analysis, Support Vector Machine, TikTok Shop, Twitter*



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	vi
HALAMAN PERSEMBAHAN/MOTO	vii
KATA PENGANTAR	viii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiv
DAFTAR KODE	xv
DAFTAR LAMPIRAN	xvi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	6
2.1 Tweet Harvest (Twitter Crawler)	6
2.2 Analisis Sentimen	6
2.3 TF-IDF (<i>Term Frequency-Inverse Document Frequency</i>)	6
2.4 Naive Bayes Classifier	7
2.4.1 Multinomial Naive Bayes	7
2.4.2 Gaussian Naive Bayes	8
2.4.3 Bernoulli Naive Bayes	8
2.5 Support Vector Machine	8
2.5.1 Kernel Linear	9
2.5.2 Kernel Polynomial	9
2.5.3 Kenel RBF	9
2.6 <i>Confusion Matrix</i>	10
2.6.1 <i>Accuracy</i>	11
2.6.2 <i>Precision</i>	11
2.6.3 <i>Recall</i>	11
2.6.4 <i>F1-Score</i>	12
BAB 3 METODOLOGI PENELITIAN	13
3.1 Metodologi Penelitian	13
3.2 Perancangan Sistem	14
3.2.1 Gambaran Umum Perancangan	14
3.2.2 Pengumpulan Data	15
3.2.3 Sorting Data	15
3.2.4 Pelabelan Data	15
3.2.5 Text Preprocessing	15
3.2.6 <i>Train-Test Split Data</i>	17
3.2.7 <i>Apply TF-IDF</i>	17

3.2.8	<i>Apply Naive Bayes</i>	17
3.2.9	<i>Apply Support Vector Machine</i>	18
3.2.10	Evaluasi	19
BAB 4	HASIL DAN DISKUSI	20
4.1	Spesifikasi Sistem	20
4.2	Implementasi Sistem	20
4.2.1	Pengumpulan Data	20
4.2.2	Sorting dan Pelabelan Data	22
4.2.3	Text Preprocessing	23
4.2.4	Training dan Testing Split Data	28
4.2.5	Apply TF-IDF	28
4.2.6	Apply Naive Bayes	29
4.2.7	Apply Support Vector Machine	29
4.2.8	Evaluasi Model	29
4.3	Uji Coba	31
4.3.1	Pengujian Menggunakan <i>Gaussian Naive Bayes</i> , <i>Multinomial Naive Bayes</i> , dan <i>Bernoulli Naive Bayes</i> pada Metode <i>Naive Bayes</i>	31
4.3.2	Pengujian Menggunakan <i>Kernel Linear</i> , <i>Polynomial</i> dan <i>RBF</i> pada Metode <i>Support Vector Machine</i>	32
4.3.3	Pengujian Pembagian Data Testing dan Data Training	32
4.3.4	Pengujian Dengan Penginputan Teks Bahasa Indonesia Secara Eksternal	33
BAB 5	SIMPULAN DAN SARAN	35
5.1	Simpulan	35
5.2	Saran	35
	DAFTAR PUSTAKA	36



DAFTAR GAMBAR

Gambar 2.1	<i>Confusion Matrix</i>	10
Gambar 3.1	Flowchart Gambaran Umum Perancangan	14
Gambar 3.2	Flowchart Text Preprocessing	16
Gambar 3.3	Flowchart Apply TF-IDF	17
Gambar 3.4	Flowchart Apply Naive Bayes	18
Gambar 3.5	Flowchart Apply Support Vector Machine	18
Gambar 4.1	Tampilan <i>Crawling Data</i> Menggunakan <i>Tweet-Harvest</i>	21
Gambar 4.2	Tampilan <i>Crawling Data</i> Menggunakan <i>Tweet-Harvest</i>	22
Gambar 4.3	Confusion Matrix Hasil Pemodelan	30
Gambar 4.4	Hasil Input Teks	34
Gambar 4.5	Hasil Input Teks	34



DAFTAR TABEL

Tabel 4.1	Tabel Pelabelan Data	22
Tabel 4.2	Tabel Hasil Proses <i>Case Folding</i>	23
Tabel 4.3	Tabel Hasil Proses <i>Cleaning Data</i>	24
Tabel 4.4	Table Hasil Proses <i>Tokenizing</i>	25
Tabel 4.5	Tabel Hasil Proses <i>Tokenizing</i>	26
Tabel 4.6	Tabel Hasil Proses <i>Stopwords</i>	27
Tabel 4.7	Table Hasil <i>Stemming</i>	28
Tabel 4.8	Metode Naive Bayes	31
Tabel 4.9	Metode Support Vector Machine	32
Tabel 4.10	Tabel Perbandingan Dataset 60:40	32
Tabel 4.11	Tabel Perbandingan Dataset 70:30	32
Tabel 4.12	Tabel Perbandingan Dataset 80:20	33



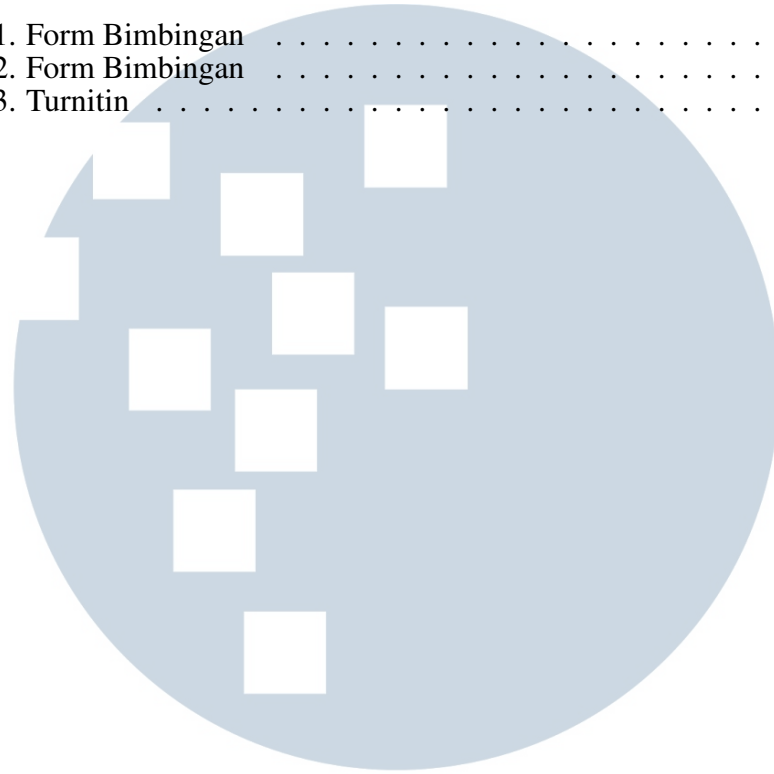
DAFTAR KODE

4.1	Potongan Kode <i>Case Folding</i>	23
4.2	Potongan Kode <i>Cleaning Data</i>	23
4.3	Potongan Kode <i>Tokenizing</i>	25
4.4	Potongan Kode <i>Normalisasi</i>	25
4.5	Potongan Kode <i>Stopword</i>	26
4.6	Potongan Kode <i>Stemming</i>	27
4.7	Potongan Kode <i>Input Teks</i>	28
4.8	Potongan Kode <i>TF-IDF</i>	28
4.9	Potongan Kode <i>Apply Naive Bayes</i>	29
4.10	Potongan Kode <i>Apply Support Vector Machine</i>	29
4.11	Potongan Kode <i>Evaluasi Model</i>	29
4.12	Potongan Kode <i>Input Teks Naive Bayes</i>	33
4.13	Potongan Kode <i>Input Teks SVM</i>	34



DAFTAR LAMPIRAN

Lampiran 1. Form Bimbingan	37
Lampiran 2. Form Bimbingan	39
Lampiran 3. Turnitin	40



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA